



099374760 .01240

#3

5 lovE Variant Regulator Molecules  
(Atty Docket No. 109272.150; Client Docket No. MIC005US)

## BACKGROUND OF THE INVENTION

### Field of the Invention

The invention relates to the fields of microbiology and molecular biology. In particular, the invention relates to the field of mycology and the production of secondary metabolites from fungi.

### Summary of the Related Art

Secondary metabolites are a major source of commercially useful products such as food additives, vitamins, and medicines for the treatment of a wide variety of infections and diseases. By way of example, in 1997 the statin drugs lovastatin, simvastatin, and pravastatin, fungal secondary metabolites used in the treatment of hypercholesterolemia, together had US sales of US\$7.53 billion (Sutherland *et al.*, *Current Opinion In Drug Discovery & Development* 4:229-236 (2001)). The cost and availability of these plant, bacterial and fungal metabolites are frequently determined by limitations imposed on production and purification of these compounds from culture. This problem is frequently exacerbated by the fact that these products are generally produced during the stationary phase of bacterial and fungal growth.

A wide variety of methods have been utilized to increase the amount of secondary metabolite produced in culture. Studies have demonstrated the importance of carefully designing the medium in which a fungus is grown to maximize the amount of a secondary metabolite produced (see, e.g., Hajjaj H, et al., *Appl. Environ. Microbiol.* 67:2596-602 (2001); Lesova, K., et al., *J. Basic Microbiol.* 40:369-75 (2000)). In addition, the method of

5 culture or fermentation also impacts directly on the amount of secondary metabolite produced. For example, see Robinson, T., et al. (*Appl. Microbiol. Biotechnol.* 55:284-289 (2001)), which demonstrates the advantages of solid state (substrate) fermentation.

10 In addition to the manipulation of culture and media conditions, genetic approaches have been taken to increase secondary metabolite production. For example, the production of penicillin is limited by the activity of two enzymes, encoded by the *ipnA* and *acvA* genes, both 15 of which are regulated by the *pacC* protein, a zinc-finger transcription factor. Naturally occurring mutant alleles of the *pacC* locus are known to possess more transcription-activating activity than the cognate, wild-type allele (see, e.g., Tilburn et al. *EMBO J.* 14(4):779-20 790 (1995)). Thus, one genetic approach to increasing secondary metabolite production is to identify and isolate naturally occurring mutant alleles, the expression of which leads to increased secondary metabolite production.

25 Although many regulators of secondary metabolite production in many organisms are known, not all of the organisms that produce secondary metabolites are amenable to genetic or molecular genetic manipulation. Thus, these systems are not generally useful as a source for 30 the isolation of naturally occurring mutant alleles and are even less useful for the deliberate manipulation of secondary metabolite regulator protein structure with the aim of creating improved regulators of secondary metabolite production.

35 It would be advantageous to have improved regulators of the biosynthetic enzymes responsible for secondary metabolite production. For example, recent studies suggest increasing usage of statin drugs, e.g., see Waters D.D., *Am. J. Cardiol.* 88:10F-5F (2001)). Thus,

- 5 demand for statin drugs is likely to increase substantially. In order to meet the demand for these and other secondary metabolites, new and improved methods for the production of secondary metabolites must be identified.

5

BRIEF SUMMARY OF THE INVENTION

The invention provides improved secondary metabolite regulator proteins that enable increased production of secondary metabolites. The invention also provides methods to make these improved regulator proteins.

10 In a first aspect, the invention provides a variant regulator protein of secondary metabolite production with increased activity than that of the cognate, wild-type protein. In certain embodiments of this aspect of the invention, the regulator protein is a fungal regulator protein.

15 In an embodiment of the first aspect, the invention provides an improved regulator protein comprising an amino acid sequence coding for a variant lovE protein having at least one specific mutation that gives rise to 20 greater transcription-activating properties of the regulator protein and/or induction of secondary metabolite synthesis.

25 By way of non-limiting example, certain preferred regulator proteins of this aspect of the invention include at least one of the following mutations: (1) a Group 6 amino acid residue mutated to a Group 2 amino acid residue at position 31, in one embodiment the mutation represented by F31L; (2) a Group 3 amino acid residue mutated to a Group 5 amino acid residue at 30 position 41, in one embodiment the mutation represented by Q41K or Q41R; (3) a Group 4 amino acid residue mutated to a Group 2 amino acid residue at position 52, in one embodiment the mutation represented by T52I; (4) a Group 4 amino acid residue mutated to a Group 3 amino acid residue at position 52, in one embodiment the mutation represented by T52N; (5) a Group 4 amino acid residue mutated to a Group 5 amino acid residue at position 73, in one embodiment the mutation represented by C73R; (6) a 35 Group 1 amino acid residue mutated to a Group 4 amino

5 acid residue at position 101, in one embodiment the mutation represented by P101S; (7) a Group 1 amino acid residue mutated to a Group 3 amino acid residue at position 101, in one embodiment the mutation represented by P101Q; (8) a valine amino acid residue mutated to another Group 2 amino acid residue at position 111, in one embodiment the mutation represented by V111I; (9) a Group 4 amino acid residue mutated to a Group 2 amino acid residue at position 133, in one embodiment the mutation represented by S133L; (10) a Group 3 amino acid residue mutated to a Group 2 amino acid residue at position 141, in one embodiment the mutation represented by E141V; (11) a Group 3 amino acid residue mutated to a Group 5 amino acid residue at position 141, in one embodiment the mutation represented by E141K; (12) a Group 4 amino acid residue mutated to Group 6 amino acid residue at position 153, in one embodiment the mutation represented by C153Y; (13) a Group 4 amino acid residue mutated to a Group 5 amino acid residue at position 153, in one embodiment the mutation represented by C153R; (14) a Group 4 amino acid residue mutated to a Group 1 amino acid residue at position 281, in one embodiment the mutation represented by T281A; (15) a Group 3 amino acid residue mutated to a Group 2 amino acid residue at position 367, in one embodiment the mutation represented by N367I; (16) a Group 3 amino acid residue mutated to a Group 6 amino acid residue at position 367, in one embodiment the mutation represented by N367Y; (17) a Group 1 amino acid residue mutated to Group 4 amino acid residue at position 389, in one embodiment the mutation represented by P389S; and (18) a Group 1 amino acid residue mutated to a Group 2 amino acid residue at position 389, in one embodiment the mutation represented by P389L.

5 In some embodiments of the first aspect, the invention provides regulator proteins with at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least eleven, or at 10 least twelve, or at least thirteen, or at least fourteen, or at least fifteen, or at least sixteen, or at least seventeen, or at least eighteen of the above described specific mutations.

15 In other embodiments of the first aspect, the invention provides an isolated lovE variant regulator protein selected from the group consisting of SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, and SEQ ID NO:65.

20 In a second aspect, the invention provides a nucleic acid molecule encoding a lovE regulator of the first aspect of the invention. By way of non-limiting example, the invention provides a nucleic acid molecule encoding the lovE variant regulator protein selected from the group consisting of SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, and SEQ ID NO:90.

25 30 35 In a third aspect, the invention provides a method of increasing the activity of a protein that regulates secondary metabolite production comprising: (a) selecting a nucleic acid comprising a polynucleotide encoding a protein regulator of secondary metabolite production; (b)

5 mutating the nucleic acid to create a plurality of nucleic acid molecules encoding variant regulator proteins of secondary metabolite production; and (c) selecting a variant regulator protein with more activity than the cognate, wild-type protein.

10 In various embodiments of the third aspect, the secondary metabolite is a fungal secondary metabolite. In certain embodiments of the third aspect, the protein regulator of secondary metabolite production is a transcription factor. In certain embodiments of the third aspect, the protein regulator of secondary metabolite production is a transmembrane transporter, protein that mediates secretion, kinase, G-protein, cell surface receptor, GTPase activating protein, guanine nucleotide exchange factor, phosphatase, protease, phosphodiesterase, bacterial protein toxin, importin, RNA-binding protein, SCF complex component, adherin, or protein encoded within a biosynthetic cluster. In certain other embodiments of the third aspect, the variant regulator protein is selected to have more activity in a heterologous cell and/or more activity in a homologous cell than the cognate, wild-type regulator protein. In certain embodiments, the variant regulator protein is selected to have more activity in a heterologous cell and/or more activity in a homologous cell than the cognate, wild-type protein and to cause more secondary metabolite to be produced in a homologous cell and/or a heterologous cell when compared to the cognate, wild-type regulator protein. In a particularly preferred embodiment, the variant regulator protein is a lovE variant regulator protein.

30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
In a fourth aspect, the invention provides a method of increasing production of a secondary metabolite comprising: (a) selecting a nucleic acid comprising a

5 polynucleotide encoding a protein regulator of secondary  
metabolite production; (b) mutating the nucleic acid to  
create a plurality of nucleic acid molecules encoding  
variant regulator proteins of secondary metabolite  
production; (c) selecting a variant regulator protein  
10 with more activity than the cognate, wild-type protein;  
and (d) expressing the selected variant regulator protein  
in a cell, thereby increasing production of the secondary  
metabolite in the cell.

In various embodiments of the fourth aspect, the  
15 secondary metabolite is a fungal secondary metabolite. In  
certain embodiments of the third aspect, the protein  
regulator of secondary metabolite production is a  
transcription factor. In certain embodiments of the  
fourth aspect, the protein regulator of secondary  
20 metabolite production is a transmembrane transporter, a  
protein that mediates secretion, a kinase, a G-protein, a  
cell surface receptor, a GTPase activating protein, a  
guanine nucleotide exchange factor, a phosphatase, a  
protease, a phosphodiesterase, a bacterial protein toxin,  
25 an importin, an RNA-binding protein, an SCF complex  
component, an adherin, or a protein encoded within a  
biosynthetic cluster. In certain other embodiments of  
the fourth aspect, the variant regulator protein is  
selected to have more activity in a heterologous cell  
30 and/or more activity in a homologous cell. In certain  
embodiments, the variant regulator protein is selected to  
have more activity in a heterologous cell and/or more  
activity in a homologous cell and to cause more secondary  
metabolite to be produced in a homologous cell and/or a  
35 heterologous cell when compared to the cognate, wild-type  
regulator protein. In a particularly preferred

5 embodiment, the variant regulator protein is a lovE variant regulator protein.

In a fifth aspect, the invention provides an isolated variant regulator protein of secondary metabolite production having increased activity compared 10 to a cognate, wild-type protein, the variant regulator protein made by the process comprising: (a) selecting a nucleic acid comprising a polynucleotide encoding a protein regulator of secondary metabolite production; (b) mutating the nucleic acid to create a plurality of 15 nucleic acid molecules encoding variant regulator proteins of secondary metabolite production; (c) selecting a variant regulator protein with more activity than the cognate, wild-type protein; and (d) recovering the selected variant regulator protein.

20 In certain embodiments of the fifth aspect, the secondary metabolite is a fungal secondary metabolite. In certain embodiments of the fifth aspect, the protein regulator of secondary metabolite production is a transcription factor. In certain embodiments of the fifth 25 aspect, the protein regulator of secondary metabolite production is a transmembrane transporter, a protein that mediates secretion, a kinase, a G-protein, a cell surface receptor, a GTPase activating protein, a guanine nucleotide exchange factor, a phosphatase, a protease, a phosphodiesterase, a bacterial protein toxin, an 30 importin, an RNA-binding protein, an SCF complex component, an adherin, or a protein encoded within a biosynthetic cluster. In certain embodiments of the fifth aspect, the variant regulator protein has more 35 activity in a heterologous and/or a homologous cell than the cognate, wild-type protein. In certain embodiments of the fourth aspect, the variant regulator protein

5 increases production of a secondary metabolite in a heterologous cell and/or a homologous cell when compared to the cognate, wild-type protein. In a particularly preferred embodiment, the variant regulator protein is a lovE variant regulator protein.

10 In a sixth aspect, the invention provides a fungus having improved lovastatin production made by the process of transforming a fungal cell with a nucleic acid molecule encoding a lovE variant protein of the first aspect of the invention. In an embodiment thereof, the 15 nucleic acid molecule is selected from a nucleic acid molecule of the second aspect of the invention.

10 In a seventh aspect, the invention provides an improved process for making lovastatin comprising transforming a fungal cell with a nucleic acid molecule 20 encoding a variant of the lovE protein of the first aspect of the invention. In an embodiment thereof, the fungal cell is transformed with a nucleic acid molecule of the second aspect of the invention.

25 In a eighth aspect, the invention provides a nucleic acid molecule encoding a lovE protein defined by SEQ ID NO:91. In an embodiment thereof, the invention provides an isolated lovE nucleic acid molecule defined by SEQ ID NO:92.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a photographic representation of cells growing on media with and without G418 selection 10 demonstrating *lovFp-HIS3p-Neo* activation in *S. cerevisiae*. Controls include MB968 (vector only), MB2478 (lowly expressed wild-type *lovE*), and MB1644 (highly expressed wild-type *lovE*). All *lovE* variants are expressed in an MB968 vector backbone similar to MB2478.

15

Figure 2A is a graphic representation of *lovFp-CYC1p-lacZ* expression in *S. cerevisiae* strains expressing *lovE* variant proteins from the clones *lovE* 1-10.

20 Figure 2B is a graphic representation of *lovFp-CYC1p-lacZ* expression in *S. cerevisiae* strains expressing *lovE* variant proteins from the clones *lovE* 1-10 from a separate transformation than that of Figure 2A.

25 Figure 3 is a graphic presentation of *lovFp-CYC1p-lacZ* expression in *S. cerevisiae* strains expressing *lovE* variant proteins from clones *lovE* 16-41.

30 Figure 4 is a graphic presentation of *lovFp-lacZ* expression in *S. cerevisiae* strains expressing *lovE* variant proteins from clones *lovE* 1-10.

35 Figure 5 is a graphic presentation of *lovFp-lacZ* expression in *S. cerevisiae* strains expressing *lovE* variant proteins from clones *lovE* 16, 20, 21, 30-34, and 36-41.

Figure 6 is a graphic presentation of lovastatin culture concentration, as measured by enzyme inhibition

5 assay, from broths of *A. terreus* cultures expressing *lovE*  
variant proteins 1-10 in.

10 Figure 7A is a graphic depiction of lovastatin  
culture concentration, as measured by HPLC analysis, from  
broths of *A. terreus* cultures expressing *lovE* variant  
15 proteins 1-10 in MF117.

Figure 7B is a graphic depiction of lovastatin  
culture concentration, as measured by HPLC analysis, from  
15 broths of *A. terreus* cultures expressing *lovE* variant  
proteins 2, 6, 30, 32, 36, 37, 39, and 41 in MF117.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The patents and publications cited herein reflect the level of knowledge in the art and are hereby incorporated by reference in their entirety. Any 10 conflict between any teaching of such references and this specification shall be resolved in favor of the latter.

The invention utilizes techniques and methods common to the fields of molecular biology, genetics and microbiology. Useful laboratory references for these 15 types of methodologies are readily available to those skilled in the art. See, for example, Molecular Cloning, A Laboratory Manual, 3<sup>rd</sup> edition, edited by Sambrook, J., MacCallum, P., and Russell, D.W. (2001), Cold Spring Harbor Laboratory Press (ISBN: 0-879-69576-5); Current 20 Protocols In Molecular Biology, edited by Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Struhl, K. (1993), John Wiley and Sons, Inc. (ISBN: 0-471-30661-4); PCR Applications: Protocols for Functional 25 Genomics, edited by Innis, M.A., Gelfand, D.H., Sninsky, J.J. (1999), Cold Spring Harbor Press (ISBN: 0-123-72186-5); and Methods In Yeast Genetics, 2000 Edition: A Cold Spring Habor Laboratory Course Manual, by Burke, D., Dawson, D. and Stearns, T., Cold Spring Harbor Press (ISBN: 0-879-69588-9).

30 In certain embodiments of the aspects of the invention, the invention relates to the biosynthesis and improved production of secondary metabolites. The invention provides variant regulator proteins useful for the production of secondary metabolites, nucleic acid 35 molecules encoding variant regulator proteins, and methods for their production.

In a first aspect, the invention provides a variant regulator protein of secondary metabolite production with increased activity relative to a cognate, wild-type

5 regulator protein. Particularly preferred are variant regulator proteins of fungal secondary metabolites.

As used herein, the terms "fungal" and "fungus" refer generally to eukaryotic, heterotrophic organisms with an absorptive mode of nutrition. Fungi typically 10 contain chitin in their cell walls and exhibit mycelial or yeast-like growth habits (More Gene Manipulations in Fungi, edited by J.W. Bennet and L.L. Lasure, Academic Press Inc. (1991), ISBN 0120886421). More specifically, the terms refer to secondary metabolite producing 15 organisms including, without limitation, *Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*, *Yarrowia lipolytica*, *Nodulisporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps* sp., *Trichoderma* sp., *Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp., *Emericellopsis* sp., 20 *Cephalosporium* sp., *Cochliobolus* sp., *Helminthosporium* sp., *Agaricus brunescens*, *Ustilago maydis*, *Neurospora* sp., *Pestalotiopsis* sp. and *Phaffia rhodozyma* (See, Fungal Physiology, Chapter 9 (Secondary(Special) Metabolism), Griffin, D. H., John Wiley & Sons, Inc.; 25 ISBN: 0471166154).

The term "variant regulator protein" is used herein to refer to any regulatory protein having at least one change or difference in the amino acid sequence of the protein when compared to its cognate, wild-type 30 regulatory protein sequence. The term does not include naturally occurring allelic variations of the cognate, wild-type regulatory protein.

The term "regulator protein" is meant to refer to a protein having a positive or negative function that 35 modifies the production of a secondary metabolite. The function of the protein may be at the level of transcription, e.g., repression or activation, protein synthesis, or transport. The regulator may alter the level of transcription, RNA stability, translation, post-

5 translational modification, or cellular localization of proteins involved in secondary metabolite synthesis and/or transport. The regulator may also have effects on precursor metabolite pools, flux through specific pathways and metabolite resistance.

10 By way of non-limiting example, certain embodiments of the aspects of the invention relate to a regulator protein that is a protein that contributes and/or promotes transcription of a gene sequence, i.e., a transcription-activating protein. "Transcription-activating" is a term used to refer to characteristics of a protein that promote transcription. As used herein, a transcription-activating protein would include proteins that increase accessibility of the DNA to transcription complexes, for example, by opening or relaxing chromatin 15 structure, proteins that promote the recognition and/or binding of transcription complexes to a target gene sequence, and/or proteins that promote transcription complex movement along the length of the template DNA 20 sequence.

25 Regulatory proteins of secondary metabolite production and the nucleic acid sequences encoding these are known to those skilled in the art. Non-limiting examples of regulatory proteins of secondary metabolite synthesis include: regulator proteins of the 30 aflatoxin/sterigmatocystin biosynthetic cluster (Woloshuk, C.P., et al., *Appl. Environ. Microbiol.* 60:2408-2414 (1994) and Brown, D.W., et al., *Proc Natl Acad Sci U S A.* 93:1418-1422 (1996)); regulator proteins of the paxilline biosynthetic cluster (Young, C., et al., 35 *Mol. Microbiol.* 39:754-764 (2001)); regulator proteins of the cephalosporin and penicillin biosynthetic clusters (Litzka O., et al., *Antonie Van Leeuwenhoek* 75:95-105 (1999); Schmitt E.K. and Kuck U., *J. Biol. Chem.* 275:9348-9357 (2000); MacCabe et al. *Mol. Gen. Genet.*

5 250:367-374 (1996); Suarez et al. *Mol. Microbiol.*  
 20:529-540 (1996); Lambert et al. *Mol. Cell. Biol.*  
 17:3966-3976 (1997); Su et al. *Genetics* 133:67-77 (1993);  
 regulator proteins of trichothecene synthesis (Trapp S.C.,  
 et al., *Mol. Gen. Genet.* 257:421-432 (1998); Brown D.W.,  
 10 et al., *Fungal Genet. Biol.* 32:121-133 (2001); and  
 Matsumoto G., et al. *Biosci. Biotechnol. Biochem.*  
 63:2001-2004 (1999)); and regulator proteins of  
 lovastatin synthesis (Kennedy, J., et al., *Science*  
 284:1368-1372 (1999); Hendrickson et al., *Chem. Biol.*  
 15 6:429-439 (1999) Tag, A. et al., *Mol Microbiol.* 38:658-65  
 (2000)).

20 Certain embodiments of the aspects of the invention disclosed herein relate to the *lovE* regulator protein, a protein which plays a key role in the biosynthesis of lovastatin. More particularly, certain embodiments of the aspects of the invention relate to variant proteins of the *lovE* regulator protein and methods of making the same. Such proteins are variant with respect to the following *A. terreus* wild-type *lovE* sequences (SEQ ID 25 NOS:91 and 92).

**Table 1: Amino Acid and Nucleic Acid Sequences of Wild-type *lovE***

**Wild-type *lovE* Amino Acid Sequence**

maadqgiftonsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrc  
 qqaglrcvysercpkrklrqsraadlvsadpdpclhmsppvpsqslpldvseshssnts  
 rqlfdppdsydwswtsigtdeaidtdcwglsqcdggfscqleptlpdlpfspfestvekap  
 lppvssdiaraasaqrelfddlsavsqeleeillavtviewpkqeiwthpigmffnasrrl  
 ltvrlrqqaqadchqgtldeclrtknlftavhcyilnvriltaiselllsqirrtqnshms  
 plegsrqspqrddtssssghssvdtipffsenlpigelfsyvdplthalfsacttlhv  
 vqlreneitlgvhshaqgiaasismsgepgediartgatnsarceeqpttpaarvlfmfl  
 sdecafqeqaksagsrgrtiaalrrcyedifslarkhkhgmlrdlnnipp (SEQ ID NO: 91)

**Wild-type *lovE* DNA Sequence**

atggctgcagatcaaggatattcacgaactcggtcactctctcgccagtggagggttca  
 cgcaccgggtggAACATTACCCGCCGTgcattccgacgctcttgtgatcggtgtcatgca  
 caaaagatcaaATGTTACTGGAATAAGGAGGTTACTGGCCGTGCTCCCTGTcagcgttgc  
 cagcaggctggacttcgatcgctcacagtgagcgatccccaaagcgcaagctacgcca  
 tccagggcagcggatctcgctctgtgaccagatccctgatggcacatgtcctcgcc  
 ccagtgcctcacagagcttgcgcgtacgttatccgagtcgcattcctcaaatacctcc  
 cgccaatttcttgcattccaccggacagctacgactggatcgatggactcgattggactgac

```

gaggctattgacactgactgctggggctgtcccaatgtgatggaggcttcagctgtcag
ttagagccaaacgctgccggatctaccttcgcgccttcgagtctacggttaaaaagctccg
ttgccaccggatcgagcgacattgtcgctgcggccagtgcgcacagagagctttcgat
gacctgtcggcggtgcgcaggaacttggaaagagatccttcgcgtacggtagaatgg
ccgaagcagggaaatctggaccatcccatcggaatgttttcaatgcgtcagcacggcctt
cttactgtcctgcgcacaagcgcaggccactgcccataaggcacactagacgaatgt
ttacggaccaagaacctcttacggcagttacactgttacatattgaatgtgcggattttg
accgcacatcgagttgcgcacattaggcgaccagaacagccatatgagc
ccacttggaaaggagtcgatcccagtcgcgcagcagacaccagcagcagcgc
cacagcagtgttgcacccataccctttagcgagaaacctccattggtagctgttc
tcctatgttgcgcctgacacacgccttatttcgcgttgcactacgttacatgttggg
gtacaattgtcgctgagaatgagattacttgcggagttacactccgcgcaggcattgca
gcttcatcagcatgagcgggaaccaggcgaggatatgcaggacaggcgcaccaat
tcgcgaagatgcgaggagcagccgaccactccagcgcgtcggttttgcattgttcttgc
agtgtgaaggggcttccaggaggcaaagtctgtcggttcccgagggtcgaaccatgcga
gcactgcgacgatgtatgaggatattttccctgcggcccaaacacaaacatggcatg
ctcagagacctaacaatattcctccatga (SEQ ID NO:92)

```

5

As used herein, the term "secondary metabolite" means a compound, derived from primary metabolites, that is produced by an organism, is not a primary metabolite, is not ethanol or a fusel alcohol, and is not required for growth under standard conditions. Secondary metabolites are derived from intermediates of many pathways of primary metabolism. These pathways include, without limitation, pathways for biosynthesis of amino acids, the shikimic acid pathway for biosynthesis of aromatic amino acids, the polyketide biosynthetic pathway from acetyl coenzyme A (CoA), the mevalonic acid pathway from acetyl CoA, and pathways for biosynthesis of polysaccharides and peptidopolysaccharides. Collectively, secondary metabolism involves all primary pathways of carbon metabolism. Particularly preferred in embodiments of the aspects of the invention are fungal secondary metabolites (See, Fungal Physiology, Chapter 9 (Secondary(Special) Metabolism), Griffin, D. H., John Wiley & Sons, Inc.; ISBN: 0471166154).

"Secondary metabolite" also includes intermediate compounds in the biosynthetic pathway for a secondary metabolite that are dedicated to the pathway for

5 synthesis of the secondary metabolite. "Dedicated to the pathway for synthesis of the secondary metabolite" means that once the intermediate is synthesized by the cell, the cell will not convert the intermediate to a primary metabolite. "Intermediate compounds" also include  
10 secondary metabolite intermediate compounds which can be converted to useful compounds by subsequent chemical conversion or subsequent biotransformation. As such, providing improved availability of such intermediate compounds would still lead to improved production of the  
15 ultimate useful compound, which itself may be referred to herein as a secondary metabolite. The yeast *Saccharomyces cerevisiae* is not known to produce secondary metabolites.

The term "primary metabolite" means a natural  
20 product that has an obvious role in the functioning of almost all organisms. Primary metabolites include, without limitation, compounds involved in the biosynthesis of lipids, carbohydrates, proteins, and nucleic acids. The term "increasing the yield of the  
25 secondary metabolite" means increasing the quantity of the secondary metabolite present in the total fermentation broth per unit volume of fermentation broth or culture.

As used herein, the phrase "modulate production of a  
30 secondary metabolite" refers to a positive or negative or desirable change in one or more of the variables or values that affect the process or results of production of the primary or secondary metabolites in a liquid or solid state fungal fermentation. These positive or  
35 negative or desirable changes include, without limitation, an increase or decrease in the amount of a primary or secondary metabolite being produced (in absolute terms or in quantity per unit volume of fermentation broth or per unit mass of solid substrate);

5 a decrease in the volume of the broth or the  
 mass/quantity of substrate required for the production of  
 sufficient quantities; a decrease in the cost of raw  
 materials and energy, the time of fermentor or culture  
 run, or the amount of waste that must be processed after  
 10 a fermentor run; an increase or decrease in the specific  
 production of the desired metabolite (both in total  
 amounts and as a fraction of all metabolites and side  
 products made by the fungus); an increase or decrease in  
 the percent of the produced secondary metabolite that can  
 15 be recovered from the fermentation broth or culture; and  
 an increase in the resistance of an organism producing a  
 primary or secondary metabolite to possible deleterious  
 effects of contact with the secondary metabolite.

In certain embodiments of aspects of the invention,  
 20 a secondary metabolite is an anti-bacterial. An "anti-  
 bacterial" is a molecule that has cytoidal or cytostatic  
 activity against some or all bacteria. Preferred anti-  
 bacterials include, without limitation,  $\beta$ -lactams.

Preferred  $\beta$ -lactams include, without limitation,  
 25 penicillins and cephalosporins and biosynthetic  
 intermediates thereof. Preferred penicillins and  
 biosynthetic intermediates include, without limitation,  
 isopenicillin N, 6-aminopenicillanic acid (6-APA),  
 penicillin G, penicillin N, and penicillin V. Preferred  
 30 cephalosporins and biosynthetic intermediates include,  
 without limitation, deacetoxycephalosporin V (DAOC V),  
 deacetoxycephalosporin C (DAOC), deacetylcephalosporin C  
 (DAC), 7-aminodeacetoxycephalosporanic acid (7-ADCA),  
 cephalosporin C, 7-B-(5-carboxy-5-oxopentanamido)-  
 35 cephalosporanic acid (keto-AD-7ACA), 7-B-(4-  
 carboxybutanamido)-cephalosporanic acid (GL-7ACA), and 7-  
 aminocephalosporanic acid (7ACA).

5 In certain embodiments of aspects of the invention, the secondary metabolite is an anti-hypercholesterolemic or a biosynthetic intermediate thereof. An "anti-hypercholesterolemic" is a drug administered to a patient diagnosed with elevated cholesterol levels for the  
10 purpose of lowering the cholesterol levels. Preferred anti-hypercholesterolemics include, without limitation, lovastatin, mevastatin, simvastatin, and pravastatin.

According to other embodiments of the invention, a secondary metabolite is an immunosuppressant or a  
15 biosynthetic intermediate thereof. An "immunosuppressant" is a molecule that reduces or eliminates an immune response in a host when the host is challenged with an immunogenic molecule, including immunogenic molecules present on transplanted organs, tissues or cells. Preferred immunosuppressants include, without limitation, members of the cyclosporin family and beauverolide L. Preferred cyclosporins include, without limitation, cyclosporin A and cyclosporin C.

In certain embodiments of aspects of the invention, the secondary metabolite is an ergot alkaloid or a biosynthetic intermediate thereof. An "ergot alkaloid" is a member of a large family of alkaloid compounds that are most often produced in the sclerotia of fungi of the genus Claviceps. An "alkaloid" is a small molecule that contains nitrogen and has basic pH characteristics. The classes of ergot alkaloids include clavine alkaloids, lysergic acids, lysergic acid amides, and ergot peptide alkaloids. Preferred ergot alkaloids include, without limitation, ergotamine, ergosine, ergocristine, ergocryptine, ergocornine, ergotaminine, ergosinine, ergocristinine, ergocryptinine, ergocorninine, ergonovine, ergometrinine, and ergoclavine.

In certain embodiments of aspects of the invention, the secondary metabolite is an inhibitor of angiogenesis

5 or a biosynthetic intermediate thereof. An "angiogenesis inhibitor" is a molecule that decreases or prevents the formation of new blood vessels. Angiogenesis inhibitors have proven effective in the treatment of several human diseases including, without limitation, cancer, 10 rheumatoid arthritis, and diabetic retinopathy. Preferred inhibitors of angiogenesis include, without limitation, fumagillin and ovalicin.

In certain embodiments of aspects of the invention, the secondary metabolite is a glucan synthase inhibitor 15 or a biosynthetic intermediate thereof. A "glucan synthase inhibitor" is a molecule that decreases or inhibits the production of 1,3- $\beta$ -D-glucan, a structural polymer of fungal cell walls. Glucan synthase inhibitors are a class of antifungal agents. Preferred glucan 20 synthase inhibitors include, without limitation, echinocandin B, pneumocandin B, aculeacin A, and papulacandin.

In certain embodiments of aspects of the invention, the secondary metabolite is a member of the gliotoxin 25 family of compounds or a biosynthetic intermediate thereof. The "gliotoxin family of compounds" are related molecules of the epipolythiodioxopiperazine class. Gliotoxins display diverse biological activities, including, without limitation, antimicrobial, antifungal, 30 antiviral, and immunomodulating activities. Preferred members of the "gliotoxin family of compounds" include, without limitation, gliotoxin and aspirochlorine.

In certain embodiments of aspects of the invention, the secondary metabolite is a fungal toxin or a 35 biosynthetic intermediate thereof. A "fungal toxin" is a compound that causes a pathological condition in a host, either plant or animal. Fungal toxins could be mycotoxins present in food products, toxins produced by

5 phytopathogens, toxins from poisonous mushrooms, or  
 toxins produced by zoopathogens. Preferred fungal toxins  
 include, without limitation, aflatoxins, patulin,  
 zearalenone, cytochalasin, griseofulvin, ergochrome,  
 cercosporin, marticin, xanthocillin, coumarins,  
 10 tricothecenes, fusidanes, sesterpenes, amatoxins,  
 malformin A, phallotoxins, pentoxin, HC toxin,  
 psilocybin, bufotenine, lysergic acid, sporodesmin,  
 pulcheriminic acid, sordarins, fumonisins, ochratoxin A,  
 and fusaric acid.

15 With some certain embodiments of aspects of the  
 invention, the secondary metabolite is a modulator of  
 cell surface receptor signaling or a biosynthetic  
 intermediate thereof. The term "cell surface receptor"  
 is as used before. Modulators of cell surface receptor  
 20 signaling might function by one of several mechanisms  
 including, without limitation, acting as agonists or  
 antagonists, sequestering a molecule that interacts with  
 a receptor such as a ligand, or stabilizing the  
 interaction of a receptor and molecule with which it  
 25 interacts. Preferred modulators of cell surface  
 signaling include, without limitation, the insulin  
 receptor agonist L-783,281 and the cholecystokinin  
 receptor antagonist asperlicin.

In certain embodiments of aspects of the invention,  
 30 the secondary metabolite is a plant growth regulator or a  
 biosynthetic intermediate thereof. A "plant growth  
 regulator" is a molecule that controls growth and  
 development of a plant by affecting processes that  
 include, without limitation, division, elongation, and  
 35 differentiation of cells. Preferred plant growth  
 regulators include, without limitation, cytokinin, auxin,  
 gibberellin, abscisic acid, and ethylene.

In certain embodiments of aspects of the invention,  
 the secondary metabolite is a pigment or a biosynthetic

5 intermediate thereof. A "pigment" is a substance that imparts a characteristic color. Preferred pigments include, without limitation, melanins and carotenoids.

10 In certain embodiments of aspects of the invention, the secondary metabolite is an insecticide or a biosynthetic intermediate thereof. An "insecticide" is a molecule that is toxic to insects. Preferred insecticides include, without limitation, nodulisporic acid.

15 In certain embodiments of aspects of the invention, the secondary metabolite is an anti-neoplastic compound or a biosynthetic intermediate thereof. An "anti-neoplastic" compound is a molecule that prevents or reduces tumor formation. Preferred anti-neoplastic compounds include, without limitation, taxol (paclitaxel) and related taxoids.

20 The phrase "increased activity" is used herein to refer to a characteristic that results in an augmentation of the inherent negative or positive function of the regulatory protein.

25 The invention provides variant regulator proteins of secondary metabolite production with increased activity and methods of producing the same. The invention further provides for the identification of specific amino acid residues that are important to the functioning of 30 secondary metabolite regulator proteins. By way of non-limiting example, variant regulator proteins of the secondary metabolite regulator lovE are presented herein.

35 As known to those skilled in the art, certain substitutions of one amino acid for another may be tolerated at one or more amino acid residues of a wild-type regulator protein absent a change in the structure, activity and/or function of the wild-type protein. Such substitutions are referred to in the art as "conservative" substitutions, and amino acids may be

5 categorized into groups that identify which amino acids  
may be substituted for another without altering the  
structure and/or function of the protein.

As used herein, the term "conservative substitution" refers to the exchange of one amino acid for another in  
10 the same conservative substitution grouping in a protein sequence. Conservative amino acid substitutions are known in the art and are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge,  
15 size, and the like. In a preferred embodiment, conservative substitutions typically include substitutions within the following groups: Group 1: glycine, alanine, and proline; Group 2: valine, isoleucine, leucine, and methionine; Group 3: aspartic  
20 acid, glutamic acid, asparagine, glutamine; Group 4: serine, threonine, and cysteine; Group 5: lysine, arginine, and histidine; Group 6: phenylalanine, tyrosine, and tryptophan. Each group provides a listing  
25 of amino acids that may be substituted in a protein sequence for any one of the other amino acids in that particular group.

As stated *supra*, there are several criteria used to establish groupings of amino acids for conservative substitution. For example, the importance of the  
30 hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte and Doolittle, *Mol. Biol.* 157:105-132 (1982)). It is known that certain amino acids may be substituted for other amino acids having a similar  
35 hydropathic index or score and still retain a similar biological activity. Amino acid hydrophilicity is also used as a criteria for the establishment of conservative amino acid groupings (see, e.g., U.S. Patent No. 4,554,101).

5       Information relating to the substitution of one amino acid for another is generally known in the art (see, e.g., Introduction to Protein Architecture : The Structural Biology of Proteins, Lesk, A.M., Oxford University Press; ISBN: 0198504748; Introduction to Protein Structure, Branden, C.-I., Tooze, J., Karolinska Institute, Stockholm, Sweden (January 15, 1999); and Protein Structure Prediction: Methods and Protocols (Methods in Molecular Biology), Webster, D.M. (Editor), August 2000, Humana Press, ISBN: 0896036375).

10      15      In one embodiment of the first aspect, the invention provides an improved regulator protein comprising an amino acid sequence coding for a variant of the lovE protein having at least one specific mutation that gives rise to greater transcription-activating properties of

20      20      the regulator protein and/or increased lovastatin synthesis.

By way of non-limiting example, certain amino acid residues and mutations thereof in the lovE regulatory protein of *A. terreus* (SEQ ID NO:91) are identified by the invention described herein. Mutations at residues 31, 41, 52, 73, 101, 111, 133, 141, 153, 281, 367, and 389 of the wild-type lovE protein of *A. terreus* have been identified as being critical for the improvement of lovE regulator protein function. Those mutations include:

25      30      31L, Q41K, Q41R, T52I, T52N, C73R, P101S, P101Q, V111I, S133L, E141V, E141K, C153Y, C153R, T281A, N367I, N367Y, P389S and P389L. Each mutation, therefore, represents a change of one conservative class of amino acids for another. For example, the mutation F31L represents a change from a Group 6 amino acid residue to a Group 2 amino acid residue at position 31 of the wild-type, lovE regulator protein.

35

Thus, by way of non-limiting example, regulator proteins of this aspect of the invention include at least

5 one of the following mutations: (1) a Group 6 amino acid residue mutated to a Group 2 amino acid residue at position 31, for example, the mutation represented by F31L; (2) a Group 3 amino acid residue mutated to a Group 5 amino acid residue at position 41, for example, the  
10 mutation represented by Q41K or Q41R; (3) a Group 4 amino acid residue mutated to a Group 2 amino acid residue at position 52, for example, the mutation represented by T52I; (4) a Group 4 amino acid residue mutated to a Group 3 amino acid residue at position 52, for example, the  
15 mutation represented by T52N; (5) a Group 4 amino acid residue mutated to a Group 5 amino acid residue at position 73, for example, the mutation represented by C73R; (6) a Group 1 amino acid residue mutated to a Group 4 amino acid residue at position 101, for example, the  
20 mutation represented by P101S; (7) a Group 1 amino acid residue mutated to a Group 3 amino acid residue at position 101, for example, the mutation represented by P101Q; (8) a valine amino acid residue mutated to another Group 2 amino acid residue at position 111, for example,  
25 the mutation represented by V111I; (9) a Group 4 amino acid residue mutated to a Group 2 amino acid residue at position 133, for example, the mutation represented by S133L; (10) a Group 3 amino acid residue mutated to a Group 2 amino acid residue at position 141, for example,  
30 the mutation represented by E141V; (11) a Group 3 amino acid residue mutated to a Group 5 amino acid residue at position 141, for example, the mutation represented by E141K; (12) a Group 4 amino acid residue mutated to Group 6 amino acid residue at position 153, for example, the  
35 mutation represented by C153Y; (13) a Group 4 amino acid residue mutated to a Group 5 amino acid residue at position 153, for example, the mutation represented by C153R; (14) a Group 4 amino acid residue mutated to a Group 1 amino acid residue at position 281, for example,

5 the mutation represented by T281A; (15) a Group 3 amino acid residue mutated to a Group 2 amino acid residue at position 367, for example, the mutation represented by N367I; (16) a Group 3 amino acid residue mutated to a Group 6 amino acid residue at position 367, for example,  
10 (17) a Group 1 amino acid residue mutated to Group 4 amino acid residue at position 389, for example, the mutation represented by P389S; and/or (18) a Group 1 amino acid residue mutated to a Group 2 amino acid residue at position 389, for example, the mutation represented by P389L.

15 In other embodiments of the first aspect, the invention provides a variant of the lovE regulator protein with at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least eleven, or at least twelve, or at least thirteen, or at least fourteen, or at least fifteen, or at least sixteen, or at least seventeen, or at least eighteen of the above described specific mutations.

20 25 In other embodiments of the first aspect, the invention provides an isolated lovE variant regulator protein having the sequence of SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, or SEQ ID NO:65.

30 35 In a second aspect, the invention provides a nucleic acid molecule encoding a variant regulator protein of secondary metabolite production of the first aspect of the invention. As used herein, the terms "nucleic acid" or "nucleic acid molecule" refer to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-

5 stranded form, and unless otherwise limited, would encompass analogs of natural nucleotides that can function in a similar manner as the naturally occurring nucleotide.

10 In one embodiment of the second aspect, the invention provides a nucleic acid molecule encoding a variant protein of the lovE regulator protein of the first aspect of the invention.

15 By way of non-limiting example, the invention provides a nucleic acid molecule encoding a lovE variant regulator protein having the sequence of SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, 20 SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, or SEQ ID NO:90.

25 Poor transformation efficiency and the lack of efficient selection systems frequently precludes the screening of large numbers of variant regulator proteins of secondary metabolites in the organism from which the regulator protein is isolated. For example, there are currently certain technical obstacles to the successful screening of large numbers of variant regulator proteins in the fungus *A. terreus*, an organism that produces the 30 secondary metabolite lovastatin.

35 The invention described herein takes advantage of the genetically tractable and experimentally amenable organism *Saccharomyces cerevisiae* for screening large numbers of variant regulator proteins of secondary metabolite production. Techniques common to the field of molecular biology are well developed in *S. cerevisiae*, and large numbers of vectors are available to assist the genetic manipulation and cloning of variant regulator proteins involved in secondary metabolite production.

5 Other genetically tractable organisms could also be used for this purpose.

In a third aspect, the invention provides a method of increasing the activity of a protein that regulates secondary metabolite production comprising: (a) selecting 10 a nucleic acid comprising a polynucleotide encoding a protein regulator of secondary metabolite production; (b) mutating the nucleic acid to create a plurality of nucleic acid molecules encoding variant regulator proteins of secondary metabolite production; and (c) 15 selecting a variant regulator protein with more activity than the cognate, wild-type protein.

As used herein, "mutating" is used to refer to the deliberate alteration of at least one nucleotide residue of a wild-type, cognate nucleic acid sequence encoding a 20 regulator protein of secondary metabolite production. A deliberate alteration or change in at least one nucleotide residue of a polynucleotide may be accomplished by any method known in the art. The mutation(s) can be made *in vivo* or *in vitro* and can 25 include random, partially random or not random, *i.e.*, directed, mutagenesis techniques.

By way of non-limiting example, *in vivo* mutagenesis can be done by placing this nucleic acid molecule in a cell with a high mutation frequency, *i.e.* a mutagenic 30 strain. By way of non-limiting example, Muhlrad *et al.* (*Yeast* 8:79-82 (1992)) have developed a rapid method for localized mutagenesis of yeast genes. As a first step, the region of interest of a gene sequence is first amplified *in vitro* under error-prone polymerase chain 35 reaction (PCR) conditions. Error-prone polymerase chain reaction (PCR) is a method of introducing amino acid changes into proteins. With this technique, mutations are deliberately introduced during the PCR reaction through the use of error-prone DNA polymerases under

5 specific reaction conditions. With the Muhlrad *et al.* procedure, the PCR product is then co-transformed with a gapped plasmid containing homology to both ends of the PCR product, resulting in *in vivo* recombination to repair the gap with the mutagenized DNA.

10 There are a variety of commercially available kits that may be used to produce mutant nucleic acid molecules by error-prone PCR (see, e.g., GeneMorph™ PCR Mutagenesis Kit (Stratagene, La Jolla, California); and Diversify™ PCR Random Mutagenesis Kit (BD Biosciences Clontech, Palo 15 Alto, CA). Thus, a plurality of variant, *i.e.*, mutated, regulator proteins of secondary metabolite production may be produced using established mutagenesis techniques.

20 As used herein, the term "activity" refers to a characteristic of the regulator protein that negatively or positively affects the biological system to bring about a modulation in secondary metabolite production. By way of non-limiting example, the activity is the transcription of downstream genes involved in the biosynthetic pathway of the secondary metabolite of 25 choice. Thus, in the present example, the phrase "more activity" refers to the property of a variant regulator protein to bring about more transcription than that effected by the cognate, wild-type regulator protein.

30 In certain embodiments of the third aspect, the selected variant regulator protein has more activity in a fungal cell than the cognate, wild-type protein. In certain embodiments of the third aspect, the protein regulator of secondary metabolite production is a transcription factor. In certain embodiments of the 35 fourth aspect, the protein regulator of secondary metabolite production is a transmembrane transporter, a protein that mediates secretion, a kinase, a G-protein, a cell surface receptor, a GTPase activating protein, a

5 guanine nucleotide exchange factor, a phosphatase, a  
protease, a phosphodiesterase, a bacterial protein toxin,  
an importin, an RNA-binding protein, an SCF complex  
component, an adherin, or a protein encoded within a  
biosynthetic cluster. . In certain other embodiments of  
10 the third aspect, the selected variant regulator protein  
has more activity in a heterologous cell than the  
cognate, wild-type protein. In certain embodiments  
thereof, the heterologous cell is an organism selected  
from the group consisting of *S. cerevisiae*, *E. coli*, *A.*  
15 *nidulans*, *Candida* sp., and *N. crassa*. In yet certain  
other embodiments of the third aspect, the selected  
variant regulator protein has more activity in a  
homologous cell than the cognate, wild-type protein. In  
certain embodiments thereof, the homologous cell is an  
20 organism selected from the group consisting of  
*Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*,  
*Yarrowia lipolytica*, *Nodulisporium* sp., *Fusarium* sp.,  
*Monascus* sp., *Claviceps* sp., *Trichoderma* sp.,  
*Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp.,  
25 *Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp.,  
*Helminthosporium* sp., *Agaricus brunescens*, *Ustilago*  
*maydis*, *Neurospora* sp., *Pestalotiopsis* sp., and *Phaffia*  
*rhodozyma*.

In certain embodiments of the third aspect, the  
30 selected variant regulator protein has more activity in a  
heterologous cell and a homologous cell than the cognate,  
wild-type protein. In certain embodiments thereof, the  
heterologous cell is an organism selected from the group  
consisting of *S. cerevisiae*, *E. coli*, *A. nidulans*,  
35 *Candida* sp., and *N. crassa* and the homologous cell is an  
organism selected from the group consisting of  
*Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*,

5 *Yarrowia lipolytica*, *Nodulisporium* sp., *Fusarium* sp.,  
*Monascus* sp., *Claviceps* sp., *Trichoderma* sp.,  
*Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp.,  
*Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp.,  
*Helminthosporium* sp., *Agaricus brunescens*, *Ustilago*  
10 *maydis*, *Neurospora* sp., *Pestalotiopsis* sp. and *Phaffia*  
*rhodozyma*.

As used herein, the phrase "heterologous cell" refers to a system for gene expression, i.e., an organism for gene expression, that is one other than the organism 15 from which the selected regulator protein of secondary metabolite production has been isolated. Preferred heterologous cells include, but are not limited to, *S. cerevisiae*, *E. coli*, *A. nidulans*, and *Candida* sp., and *N. crassa*. Particularly preferred are fungal 20 heterologous cells. In an embodiment of the third aspect, the method comprises: (a) selecting a nucleic acid comprising a polynucleotide encoding a protein regulator of secondary metabolite production; (b) mutating the nucleic acid to create a plurality of 25 nucleic acid molecules encoding variant regulator proteins of secondary metabolite production; and (c) selecting a mutagenized nucleic acid encoding a variant regulator protein with increased activity in a homologous cell than the cognate, wild-type protein.

30 As used herein, the phrase "homologous cell" refers to a system for gene expression, i.e., an organism for gene expression, that is the organism from which the regulator protein of secondary metabolite production has been isolated. Preferred homologous cells are fungal 35 homologous cells, including, but not limited to, *Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*, *Yarrowia lipolytica*, *Nodulisporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps* sp., *Trichoderma* sp.,

5 *Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp.,  
*Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp.,  
*Helminthosporium* sp., *Agaricus brunescens*, *Ustilago*  
*maydis*, *Neurospora* sp., *Pestalotiopsis* sp and *Phaffia*  
*rhodozyma*. (See, Fungal Physiology, Chapter 9  
10 (Secondary (Special) Metabolism), Griffin, D. H., John  
Wiley & Sons, Inc.; ISBN: 0471166154).

In certain embodiments of the third aspect, the method further comprises selecting a variant regulator protein that also increases production of a secondary 15 metabolite in a cell when compared to the cognate, wild-type protein. In certain embodiments thereof, the cell is a fungal cell. In certain embodiments thereof, the cell is a heterologous cell, preferably selected from the group consisting of *S. cerevisiae*, *E. coli*, *A. nidulans*,  
20 *Candida* sp., and *N. crassa*.

In certain embodiments thereof, the cell is a homologous cell, preferably selected from the group consisting of *Aspergillus* sp., *Penicillium* sp.,  
*Acremonium chrysogenum*, *Yarrowia lipolytica*,  
25 *Nodulisporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps* sp., *Trichoderma* sp., *Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp., *Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp., *Helminthosporium* sp., *Agaricus brunescens*, *Ustilago maydis*, *Neurospora* sp.,  
30 *Pestalotiopsis* sp., and *Phaffia rhodozyma*.

Certain embodiments of the aspects of the invention relate to regulator proteins that promote secondary metabolite production by increasing transcription of one or more genes involved with secondary metabolite 35 production. These wild-type sequences may be selected for mutagenesis to create a plurality of variant regulator proteins. The activity of these transcription-

5 activating variant regulator proteins may be determined by measuring the activity of a reporter gene having the appropriate promoter sequences. These tests are done in a homologous and/or a heterologous cell. Certain  
10 embodiments of aspects of the invention are directed to fungal regulator proteins with transcription-activating activity that is tested in fungal heterologous and homologous cells.

Reporter genes are useful for isolating transformants expressing improved variant regulator 15 proteins. The reporter genes may be operably linked to a promoter sequence that is normally regulated by the wild-type regulator protein. Reporter genes include, but are not limited to, genes encoding  $\beta$ -galactosidase (*lacZ*),  $\beta$ -glucuronidase (*GUS*),  $\beta$ -glucosidase, amylase and  
20 invertase, amino acid biosynthetic genes, e.g., the yeast *LEU2*, *HIS3*, *LYS2*, *TRP1* genes (or homologous genes from other fungi, such as filamentous fungi, that encode proteins with the similar functional activities), nucleic acid biosynthetic genes, e.g., the yeast *URA3* and *ADE2*  
25 genes (or homologous genes from other fungi, such as filamentous fungi, that encode proteins with the similar functional activities), the mammalian chloramphenicol transacetylase (CAT) gene, or any surface antigen gene for which specific antibodies are available. A reporter  
30 gene can also be a neomycin phosphotransferase (neo) gene, which encodes neomycin, kanamycin resistance gene and G418 (geneticin) resistance gene. A reporter gene may encode a protein detectable by luminescence or fluorescence, such as green fluorescent protein (GFP).  
35 Reporter genes may additionally or alternatively encode any protein that provides a phenotypic marker, for example, a protein that is necessary for cell growth or viability, or a toxic protein that causes cell death.

5 Alternatively, the reporter gene may encode a protein detectable by a color assay leading to the presence or absence of color.

The choice of reporter gene will depend on the type of cell to be transformed. Preferred reporter genes are 10 those that are operable in fungal cells. It is preferable to have two reporter genes within the cell. One reporter gene, when expressed, provides a growth advantage to transformed cells that are expressing the variant regulator protein. This allows for the isolation 15 of such transformants though selective pressures. The other reporter gene provides a colorimetric marker, such as the *lacZ* gene and its encoded protein,  $\beta$ -galactosidase. Alternatively, the second reporter provides a fluorescent or luminescent marker, such as 20 green fluorescent protein (GFP).

In a fourth aspect, the invention provides a method of increasing production of a secondary metabolite comprising: (a) selecting a nucleic acid comprising a polynucleotide encoding a protein regulator of secondary 25 metabolite production; (b) mutating the nucleic acid to create a plurality of nucleic acid molecules encoding variant regulator proteins of secondary metabolite production; (c) selecting a variant regulator protein with more activity than the cognate, wild-type protein; 30 and (d) expressing the selected variant regulator protein in a cell, thereby increasing production of the secondary metabolite in the cell.

In certain embodiments of the fourth aspect, the cell is a fungal cell. In certain embodiments of the 35 fourth aspect, the protein regulator of secondary metabolite production is a transcription factor. In certain embodiments of the fourth aspect, the protein regulator of secondary metabolite production is a

5 transmembrane transporter, a protein that mediates secretion, a kinase, a G-protein, a cell surface receptor, a GTPase activating protein, a guanine nucleotide exchange factor, a phosphatase, a protease, a phosphodiesterase, a bacterial protein toxin, an  
10 importin, an RNA-binding protein, an SCF complex component, an adherin, or a protein encoded within a biosynthetic cluster. In certain embodiments of the fourth aspect, the cell is a heterologous cell, preferably selected from the group consisting of *S. cerevisiae*, *E. coli*, *A. nidulans*, *Candida* sp., and *N. crassa*. In certain other embodiments of the fourth aspect, the cell is a homologous cell, preferably selected from the group consisting of *Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*, *Yarrowia lipolytica*, *Noduliporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps* sp., *Trichoderma* sp., *Tolypocladium* sp., *Tricotheicum* sp., *Fusidium* sp., *Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp., *Helminthosporium* sp., *Agaricus brunescens*, *Ustilago maydis*, *Neurospora* sp., *Pestalotiopsis* sp., and *Phaffia rhodozyma*.

In certain other embodiments of the fourth aspect, the cell is a heterologous cell and the method further comprises expressing the variant regulator protein in a homologous cell, thereby increasing secondary metabolite production in the homologous cell. In certain embodiments thereof, the heterologous cell is an organism selected from the group consisting of *S. cerevisiae*, *E. coli*, *A. nidulans*, *Candida* sp., and *N. crassa* and the homologous cell is an organism selected from the group consisting of *Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*, *Yarrowia lipolytica*, *Noduliporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps*

5       *sp.*, *Trichoderma* *sp.*, *Tolypocladium* *sp.*, *Tricotheicium* *sp.*, *Fusidium* *sp.*, *Emericellopsis* *sp.*, *Cephalosporium* *sp.*, *Cochliobolus* *sp.*, *Helminthosporium* *sp.*, *Agaricus brunescens*, *Ustilago maydis*, *Neurospora* *sp.*, *Pestalotiopsis* *sp.* and *Phaffia rhodozyma*.

10       In a fifth aspect, the invention provides an isolated variant regulator protein of secondary metabolite production having increased activity compared to a cognate, wild-type protein, made by the process comprising: (a) selecting a nucleic acid comprising a polynucleotide encoding a protein regulator of secondary metabolite production; (b) mutating the nucleic acid to create a plurality of nucleic acid molecules encoding variant regulator proteins of secondary metabolite production; (c) selecting a variant regulator protein with more activity than the cognate, wild-type protein; and (d) recovering the selected variant regulator protein.

15       In certain embodiments of the fifth aspect, the variant regulator protein selected has more activity in a fungal cell. In certain embodiments of the fifth aspect, the protein regulator of secondary metabolite production is a transcription factor. In certain embodiments of the fifth aspect, the protein regulator of secondary metabolite production is a transmembrane transporter, a protein that mediates secretion, a kinase, a G-protein, a cell surface receptor, a GTPase activating protein, a guanine nucleotide exchange factor, a phosphatase, a protease, a phosphodiesterase, a bacterial protein toxin, an importin, an RNA-binding protein, an SCF complex component, an adherin, or a protein encoded within a biosynthetic cluster. In certain embodiments of the fifth aspect, the variant regulator protein selected has

5 more activity in a heterologous cell, preferably selected from the group consisting of *S. cerevisiae*, *E. coli*, *A. nidulans*, *Candida* sp., *Neurospora* sp., *Pestalotiopsis* sp., and *N. crassa*. In certain embodiments of the fifth aspect, the variant regulator protein selected has more  
10 activity in a homologous cell, preferably selected from the group consisting of *Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*, *Yarrowia lipolytica*, *Nodulisporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps* sp., *Trichoderma* sp., *Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp., *Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp., *Helminthosporium* sp., *Agaricus brunescens*, *Ustilago maydis*, *Neurospora* sp., *Pestalotiopsis* sp., and *Phaffia rhodozyma*.

In certain embodiments of the fifth aspect, the variant regulator protein selected has more activity in a homologous cell and a heterologous cell. In embodiments thereof, the heterologous cell is an organism selected from the group consisting of *S. cerevisiae*, *E. coli*, *A. nidulans*, *Candida* sp., *Neurospora* sp., *Pestalotiopsis* sp., and *N. crassa* and the homologous cell is an organism selected from the group consisting of *Aspergillus* sp., *Penicillium* sp., *Acremonium chrysogenum*, *Yarrowia lipolytica*, *Nodulisporium* sp., *Fusarium* sp., *Monascus* sp., *Claviceps* sp., *Trichoderma* sp., *Tolypocladium* sp., *Tricotheicium* sp., *Fusidium* sp., *Emericellopsis* sp., *Cephalosporium* sp., *Cochliobolus* sp., *Helminthosporium* sp., *Agaricus brunescens*, *Ustilago maydis*, *Neurospora* sp., *Pestalotiopsis* sp., and *Phaffia rhodozyma*.

In yet another embodiment of the fifth aspect, the variant regulator protein is a variant protein of the lovE protein having at least one of the following mutations: (1) a Group 6 amino acid residue mutated to a

5 Group 2 amino acid residue at position 31, for example, the mutation represented by F31L; (2) a Group 3 amino acid residue mutated to a Group 5 amino acid residue at position 41, for example, the mutation represented by Q41K or Q41R; (3) a Group 4 amino acid residue mutated to 10 a Group 2 amino acid residue at position 52, for example, the mutation represented by T52I; (4) a Group 4 amino acid residue mutated to a Group 3 amino acid residue at position 52, for example, the mutation represented by T52N; (5) a Group 4 amino acid residue mutated to a Group 15 5 amino acid residue at position 73, for example, the mutation represented by C73R; (6) a Group 1 amino acid residue mutated to a Group 4 amino acid residue at position 101, for example, the mutation represented by P101S; (7) a Group 1 amino acid residue mutated to a 20 Group 3 amino acid residue at position 101, for example, the mutation represented by P101Q; (8) a valine amino acid residue mutated to another Group 2 amino acid residue at position 111, for example, the mutation represented by V111I; (9) a Group 4 amino acid residue mutated to a Group 2 amino acid residue at position 133, for example, the mutation represented by S133L; (10) a 25 Group 3 amino acid residue mutated to a Group 2 amino acid residue at position 141, for example, the mutation represented by E141V; (11) a Group 3 amino acid residue mutated to a Group 5 amino acid residue at position 141, for example, the mutation represented by E141K; (12) a 30 Group 4 amino acid residue mutated to Group 6 amino acid residue at position 153, for example, the mutation represented by C153Y; (13) a Group 4 amino acid residue mutated to a Group 5 amino acid residue at position 153, for example, the mutation represented by C153R; (14) a 35 Group 4 amino acid residue mutated to a Group 1 amino acid residue at position 281, for example, the mutation represented by T281A; (15) a Group 3 amino acid residue

5 mutated to a Group 2 amino acid residue at position 367, for example, the mutation represented by N367I; (16) a Group 3 amino acid residue mutated to a Group 6 amino acid residue at position 367, for example, the mutation represented by N367Y; (17) a Group 1 amino acid residue 10 mutated to Group 4 amino acid residue at position 389, for example, the mutation represented by P389S; and/or (18) a Group 1 amino acid residue mutated to a Group 2 amino acid residue at position 389, for example, the mutation represented by P389L.

15 In certain embodiments of this aspect of the invention, the variant protein of the lovE protein sequence has an amino acid sequence of SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ 20 ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, or SEQ ID NO:65.

25 In another embodiment thereof, the variant protein of the lovE protein is encoded by a nucleic acid molecule having a polynucleotide sequence of SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:79, SEQ ID 30 NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, or SEQ ID NO:90.

35 In a sixth aspect, the invention provides a fungus having improved lovastatin production made by the process of transforming a fungal cell with a nucleic acid molecule encoding a variant of the lovE protein of the first aspect of the invention. In an embodiment thereof, the nucleic acid molecule is selected from a nucleic acid molecule of the second aspect of the invention.

5        In a seventh aspect, the invention provides an improved process for making lovastatin comprising transforming a fungal cell with a nucleic acid molecule encoding a variant of the *lovE* protein of the first aspect of the invention. In an embodiment thereof, the  
 10      fungal cell is transformed with a nucleic acid molecule of the second aspect of the invention.

International Patent Application PCT/US99/29583 discloses lovastatin production genes. However, this reference does not provide a mature *lovE* cDNA sequence.  
 15      The invention herein remedies the shortcoming of this reference by providing a complete cDNA sequence for the *lovE* mRNA.

In an eighth aspect, the invention provides a nucleic acid molecule encoding a *lovE* protein defined by  
 20      SEQ ID NO:91. In an embodiment thereof, the invention provides an isolated *lovE* nucleic acid molecule defined by SEQ ID NO:92.

The following examples illustrate the preferred modes of making and practicing the present invention but  
 25      are not meant to limit the scope of the invention since alternative methods may be utilized to obtain similar results.

#### EXAMPLES

30

##### **Example 1: Preparation of Strains and Plasmids**

Strain MY2124 was derived from the Sigma 1278b strain background of *S. cerevisiae* and its complete genotype is as follows: *MATα/MATα::LEU2    ura3Δ0 /ura3Δ0*  
 35      *leu2Δ0/leu2Δ0 trp1Δ0::hisG/trp1Δ0::hisG    his3Δ0::hisG/his3Δ0::hisG    ura3Δ0::lovF-HIS3p-neo/ura3Δ0.*  
*MY2124 can be constructed by mating *S. cerevisiae* strains MY2112 (MATα ura3Δ0 leu2Δ0 trp1Δ0::hisG his3Δ0::hisG*

5      *ura3Δ0::lovFp-HIS3p-neo*) with MY1555 (*matα::LEU2 ura3Δ0*  
*leu2Δ0 trp1Δ0::hisG his3Δ0::hisG*) and isolating zygotes.  
 The *ura3Δ0::lovFp-HIS3p-neo* allele of MY2112 was derived  
 by cotransforming *SfiI*-linearized plasmid MB2254 with  
 pRS424 (Sikorski and Hieter (1989) *Genetics* 122:19-27)  
 10     into MY1413 (*MATα leu2Δ0 trp1Δ0::hisG his3Δ0::hisG*).  
 Transformants were selected on SC-Trp media and  
 subsequently screened for 5-fluoro-orotic acid resistance  
 to identify those transformants containing the  
*ura3Δ0::lovFp-HIS3p-neo* allele. *Trp*<sup>r</sup> segregants lacking  
 15     plasmid pRS424 were isolated by growing the strain under  
 non-selective conditions.

The following oligonucleotides were used in the  
 construction of plasmids.

<b>Table 2: Oligonucleotides Utilized For LovE Variant Cloning</b>
<b>MO664</b> (5' GGCCATGGAGGCCGCTAGCTCGAGTCGACGGCCTAGGTGCCAGCT3' ) (SEQ ID NO:1)
<b>MO665</b> (5' GGCCACCTAGGCCGTCGACTCGAGCTAGCGGCCTCCATGGCCGTAC3' ) (SEQ ID NO:2)
<b>MO666</b> (5' GGCAGCCGCTCTAGAACTAGTCTCGAGGGTACC3' ) (SEQ ID NO:3)
<b>MO667</b> (5' GGTACCCCTCGAGACTAGTTCTAGAGCGGCCGCC3' ) (SEQ ID NO:4)
<b>MO1794</b> (5' CACAGCGGCCGCTAACCTTCCCATTGGGGC3' ) (SEQ ID NO:5)
<b>MO1793</b> (5' CACCACTAGTACGCCGGCTGATTGAC3' ) (SEQ ID NO:6)
<b>MO1785</b> (5' CACCACTAGTTATACATTATATAAGTAATGTG3' ) (SEQ ID NO:7)
<b>MO1786</b> (5' CACAGGATCCGTATCTTGCCTTCGTTATC3' ) (SEQ ID NO:8)
<b>MO195</b> (5' CGCGGATCCTATTGAACAAAGATGGATTGCAC3' ) (SEQ ID NO:9)
<b>MO196</b> (5' CCGGAATTCAAGAAGAACTCGTCAAGAAG3' ) (SEQ ID NO:10)
<b>MO841</b> (5' ACAAAAAAGCAGGCTCCACAATGGCTGCAGATCAAGGTAT3' ) (SEQ ID NO:11)
<b>MO842</b> (5' ACAAGAAAGCTGGGTTCATGGAGGAATATTGTTGA3' ) (SEQ ID NO:12)
<b>MO2278</b> (5' GGGGATCCAATCGAGGTCCACGACCAGT3' ) (SEQ ID NO:13)
<b>MO343</b> (5' GGGGACAAGTTGTACAAAAAAGCAGGCT3' ) (SEQ ID NO:14)
<b>MO2273</b> (5' GGGGATCCGCAATGGTCCCGTTCAAAC3' ) (SEQ ID NO:15)
<b>MO2274</b> (5' ACAAGAAAGCTGGGTTCACAGAAATGTTAGCTCAA3' ) (SEQ ID NO:16)
<b>MO344</b> (5' GGGGACCACTTGTACAAGAAAGCTGGGT3' ) (SEQ ID NO:17)
<b>MO2624</b> (5' GCGATGCCCAAGCGCAAGCTACGCCAATCCAGGG3' ) (SEQ ID NO:18)
<b>MO2654</b> (5' CGTCGCGCCATTGCCATTAGGCTGCGCAACTGT3' ) (SEQ ID NO:19)

<b>MO2680</b>	(5'GGACCTTGAGCATAAAATTACTATACTTCT3')	(SEQ ID NO:20)
<b>MO2686</b>	(5'GGCGCGTCCATTGCCATTAGGCTGCGCAACTGT3')	(SEQ ID NO:21)
<b>MO2681</b>	(5'TAAAACCTTGTCTTCTTCTCTAAAT3')	(SEQ ID NO:22)
<b>MO2700</b>	(5'CAGTGAGCGCGCGTAATACGACTCACTATAGGGCGA3')	(SEQ ID NO:23)
<b>MO2701</b>	(5' ATACTTCTATAGACACACAAACACAAATACACACAC3')	(SEQ ID NO:24)
<b>MO107</b>	(5'CGCGGATCCCGTCGTTTACAAC3')	(SEQ ID NO:25)
<b>MO197</b>	(5'CCCAAGCTTATTATTTGACACCCAGACCAA3')	(SEQ ID NO:26)
<b>MO1293</b>	(5'GGAAGATCTAGCATCGTGGCCAATTCTTCTAGTTT3')	(SEQ ID NO:27)
<b>MO1294</b>	(5'ATAAGAATGCGGCCGCTAACCTCCCATTGGGGCGTTGC3')	(SEQ ID NO:28)
<b>MO1787</b>	(5'CACAGGATCCAGCATTATAATTTAGTGTGTATTT3')	(SEQ ID NO:29)
<b>MO1788</b>	(5'CACCACTAGTCTCGAGCAGATCCGCCAG3')	(SEQ ID NO:30)
<b>MO1793</b>	(5'CACCACTAGTACCGGGCTGATTGAC3')	(SEQ ID NO:31)
<b>MO1794</b>	(5'CACAGCGGCCGCTAACCTCCCATTGGGGC3')	(SEQ ID NO:32)
<b>MO511</b>	(5'GGCCATCGATAAGTTGTACAAAAAGCTGAAC3')	(SEQ ID NO:33)
<b>MO540</b>	(5'GGCGCCCTATTACACCACTTGTACAAGAAAGC3')	(SEQ ID NO:34)
<b>MO1985</b>	(5'CACACGTCTCCGGCCTAACCTCCCATTGGGGCG3')	(SEQ ID NO:35)
<b>MO1986</b>	(5'CACACAGATCTCGTGGCCAATTCTTCTAGTTGA3')	(SEQ ID NO:36)
<b>MO1992</b>	(5'CACACGGATCCACAATGTTACGTCTGTAGAAACCCC3')	(SEQ ID NO:37)
<b>MO1993</b>	(5'CACAGCGGCCGCTTCATTGTTGCCTCCCTGCTG3')	(SEQ ID NO:38)
<b>MO316</b>	(5'GCGGCCGCGGCCGCCATGTCAACAAGAAT3')	(SEQ ID NO:39)
<b>MO318</b>	(5'CCGGCCGAGTGGAGATGTGGAGT3')	(SEQ ID NO:40)

5

Plasmid MB2254 contains the *lovFp-HIS3p-neo* reporter gene flanked by *URA3* sequence. First primers MO664 (SEQ ID NO:1) and MO665 (SEQ ID NO:2) were annealed and 10 inserted into the *KpnI-SacI* sites of plasmid pBluescript II KS (Stratagene,). The resulting vector, MB1038, contains a *SalI* site in the polylinker. Next, the *SpeI-XhoI* fragment from pJL164 (Brachmann et al. Yeast 14:115-132 (1998)) containing a deletion of the *URA3* gene with 15 additional flanking sequences was inserted into the *NheI-SalI* sites of MB1038 to create MB1053. Primers MO666

5 (SEQ ID NO:3) and MO667 (SEQ ID NO:4) that contain  
multiple restriction sites (*NotI*, *XbaI*, *SpeI*, *XhoI* and  
*KpnI*) were then annealed together and ligated into the  
*SmaI* site of MB1053 to create MB1054. Next, the  
10 following four fragments were combined in MB1054 to  
obtain plasmid MB2254. The *lovF* promoter from *A. terreus*  
genomic DNA was PCR amplified with MO1794 (SEQ ID NO:5)  
and MO1793 (SEQ ID NO:6) and inserted into MB1054 on a  
15 *NotI-SpeI* fragment. The *HIS3* basal promoter from pRS403  
(Sikorski and Hieter, *Genetics* 122:19-27 (1989)) was PCR  
amplified with primers MO1785 (SEQ ID NO:7) and MO1786  
(SEQ ID NO:8) and inserted into MB1054 on a *SpeI-BamHI*  
fragment. Finally, the *neo* gene (PCR amplified with  
MO195 (*BamHI*) (SEQ ID NO:) and MO196 (*EcoRI*) (SEQ ID  
NO:10) from plasmid pYX11 (Xiao and Weaver, *Nucl. Acids  
20 Res.* 25:2985-2991 (1997)) and *CYC1* terminator sequences  
(*XhoI-KpnI* fragment from pRS426-GAL-S (Mumberg, et al.,  
*Nucl. Acids. Res.* 22:5767-5768 (1994)) were first  
combined in pRS416 (Sikorski and Hieter, *Genetics* 122:19-  
27 (1989)) and then cut out with *BamHI-KpnI* and inserted  
25 into MB1054 to create MB2254.

The *lovFp-HIS3p-neo* reporter in MY2124 can confer  
resistance to the drug geneticin (G418). It was  
empirically determined that MY2124 (untransformed or  
transformed with parental plasmids MB2478 (*CYC1-lovE/CEN*)  
30 or MB2848 (*CYC1-lovE/At274/CEN*) was unable to grow on YPD  
media supplemented with 100 µg /ml G418. Plasmid MB2478  
contains the *CYC1* promoter operationally linked to the  
entire *A. terreus* *lovE* open reading frame. The *CYC1*  
promoter is a relatively weak promoter and thus the *lovE*  
35 ORF in MB2478 was expressed at low levels. MB2478 was  
the parental vector plasmid for creating full length *lovE*  
variants. Plasmid MB2848 contains the *CYC1* promoter  
operationally linked to a chimeric open reading frame

5 consisting of the *A. terreus* *lovE* DNA binding domain fused to the carboxy-terminal portion of the *At274* gene (U.S. Serial No. 60/257,431, filed December 22, 2000).

MB2848 was used to create *lovE* variants in which the DNA binding domain was not mutated. Both MB2478 and 10 MB2848 contain yeast CEN and autonomously replicating sequences and both are maintained at 1-2 copies per cell. In contrast to strains transformed with MB2478 or MB2848, strains transformed with plasmid MB1644 (*TEF1-lovE/2* micron) were able to grow on G418-supplemented YPD media. 15 The *lovE* gene of MB1644 is under control of the constitutively strong *S. cerevisiae* *TEF1* promoter. MB1644 contains a 2-micron origin for high-copy replication in yeast. An objective of these studies was to identify *lovE* variants which when expressed at low 20 levels could confer G418 resistance similar to the highly expressed wild-type *lovE* molecule of MB1644. *S. cerevisiae* expression vectors used in these studies were constructed as follows.

MB968 is a low copy *S. cerevisiae* *URA3* based 25 expression vector. MB968 was created by inserting the *EcoRV* fragment (containing the destination cassette) from gateway pEZC7201 (Invitrogen™, Carlsbad, CA) into *XhoI/SalI* (filled in with Klenow) linearized pRS416 CYC1 (Mumberg, et al., *Gene* 156:119-122 (1995)).

30 MB1644 and MB2478 are *URA3*-based *S. cerevisiae* expression plasmids that contain the wild-type *lovE* gene. They are both derivatives of MB1199. MB1199 was created by using primers M0841 (SEQ ID NO:11) and M0842 (SEQ ID NO:12) to amplify the *lovE* ORF from *A. terreus* cDNA. 35 Gateway (Invitrogen™, Carlsbad, CA) Cloning Technology (US Patent 5,888,732) was used to clone the *lovE* PCR fragment into the gateway entry vector pDONR206 (Invitrogen™, Carlsbad, CA) to create MB1199. Similarly, Gateway Cloning Technology was used to transfer the *lovE*

5 ORF from MB1199 into MB968 to create MB2478 and into  
MB969 (U.S. Serial No. 60/198,335, filed April 18, 2000)  
to create MB1644.

MB2848 is a derivative of MB968 that contains a  
lovE-AT274 chimera. The *lovE* portion of MB2848 was  
10 derived by using oligos MO841 (SEQ ID NO:11) and MO2278  
(SEQ ID NO:13) to PCR amplify the *lovE* DNA binding domain  
from *A. terreus* cDNA. A second round of PCR was  
performed with primers MO343 (SEQ ID NO:14) and MO2278 to  
add appropriate Gateway Cloning Technology compatible  
15 sequences. The *At274* portion of MB2848 can be derived by  
using primers MO2273 (SEQ ID NO:15) and MO2274 (SEQ ID  
NO:16) to PCR amplify the carboxy-terminal domain of  
*At274* from *A. terreus* cDNA. A second round of PCR was  
performed with primers MO344 (SEQ ID NO:17) and MO2273 to  
20 add appropriate Gateway Cloning Technology compatible  
sequences. The *lovE* and *At274* PCR products were cut with  
*Bam*HI and purified over a QIAquick PCR purification kit  
(Qiagen, Valencia, CA) according to manufacturer's  
instructions. Finally, the products were mixed 3-4 hours  
25 in a standard ligation reaction and used in Gateway entry  
and destination reactions to create MB2848.

Gateway cloning technology was used to clone the  
*lovE* variants of interest into plasmid MB1419 which is a  
filamentous fungal expression vector. The MB1419 fungal  
30 selection marker is the *A. nidulans* *GPD* promoter  
controlling the *ble* gene from *S. hindustanus*. The  
transgene is controlled by the *A. nidulans* *PGK* promoter.  
*A. terreus* strain MF117 is a derivative of *A. terreus*  
strain ATCC 20542.

35

**Example 2: PCR Mutagenesis of the *lovE* DNA Binding Domain**

The zinc finger DNA binding domain of *lovE* is encoded  
by nucleotides 100-201 (SEQ ID NO:92). Oligos MO2624

5 (SEQ ID NO:18) and MO2654 (SEQ ID NO:19) were used to PCR  
amplify a *lovE* containing fragment from plasmid MB2478.  
The 1.7 kb product contains nucleotides 212-1410 of *lovE*  
and ~500 bp of flanking vector sequence. Two rounds of  
standard PCR (1.5 mM MgCl<sub>2</sub>) were performed with AmpliTaq  
10 DNA polymerase (Applied Biosystems, Foster City, Ca)  
according to the manufacturer's instructions.

15 Plasmid MB2848 was cut with *Kpn*I-*Bam*HI to release a 1.1  
kb fragment containing the At274 portion of the *lovE*-  
At274 chimeric open reading frame. The remaining 5.5 kb  
vector sequence retains the *lovE* DNA binding domain.

**Example 3: PCR Mutagenesis of the *lovE* Open Reading Frame**

lovE open reading frame insert was prepared  
according to the following procedure. Oligo pairs MO2680  
20 (SEQ ID NO:20) /MO2686 (SEQ ID NO:21), MO2681 (SEQ ID  
NO:22) /MO2686, and MO2700 (SEQ ID NO:23) /MO2701 (SEQ ID  
NO:24) were used to PCR amplify the entire *lovE* open  
reading frame from plasmid MB2478. The PCR products  
differ in the amount of 5' and 3' vector sequence  
25 flanking the *lovE* open reading frame.

PCR was performed using a GeneMorph PCR mutagenesis  
kit (Stratagene, La Jolla, Ca) according to  
manufacturer's instructions to achieve medium and high  
range mutation frequencies.

30 Plasmid MB2478 was cut with *Asp*718/*Xba*I to release a  
1.7 kb fragment. The remaining 5.0 kb vector sequence  
completely lacks *lovE* ORF sequence.

**Example 4: Transformation and Selection for G418R**

35 **Isolates**

All PCR products were purified using a QIAquick PCR  
purification kit (Qiagen) according to manufacturer's  
instructions. All vectors were gel purified using a

5 QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions.

The mutagenesis strategy of Muhlrad et al. (Yeast 8:79-82 (1992)) was used which involves cotransforming a mutated PCR product and gapped plasmids into *S.*

10 *cerevisiae*, and then screening for *in vivo* recombinants having the desired phenotype).

Transformation of *Saccharomyces cerevisiae* was accomplished by the lithium acetate/single-stranded carrier DNA/polyethylene glycol (LiAc/ss-DNA/PEG) 15 protocol (Woods R.A. and Gietz R.D. *Methods Mol. Biol.* 177:85-97 (2001)) with a 1:5 molar ratio of vector:insert DNA to generate >55,000 *in vivo* recombinant transformants on SC-Ura plates. Transformants were transferred by replica printing to YPD plates containing 100 µg/ml G418 20 and allowed to grow for 2-4 days at 30°C (Figure 1).

Drug resistant clones were confirmed in secondary assays including growth on G418 concentrations up to 2000 µg/ml. The plasmid-dependence of the phenotype was determined by observing the re-appearance of drug 25 sensitivity correlating with loss of the library plasmid. *lovE* variant plasmids were recovered from promising candidates (Hoffman and Winston (1986) *Gene* 57:267). More than 70 *lovE* variants were identified and 30 definitively characterized by DNA sequence and/or restriction digestion analysis.

Table 3 summarizes the G418 resistance phenotype and sequence analysis of 26 of these variants.

Table 3: Variant loxE Mutations

<u>loxE allele</u>	<u>loxFp-neo Mediated G418R</u>	<u>MO oligos used for random PCR mutagenesis</u>	<u>Amino Acid Change 1</u>	<u>Amino Acid Change 2</u>	<u>Amino Acid Change 3</u>	<u>Amino Acid Change 4</u>	<u>Amino Acid Change 5</u>	<u>Amino Acid Change 6</u>	<u>Amino Acid Change 7</u>	<u>Amino Acid Change 8</u>	<u>Amino Acid Change 9</u>	<u>Amino Acid Change 10</u>	<u>Amino Acid Change 11</u>
1	-/+	2624/2654	H253R	R121W	S341P								
2	+/-	2624/2654		S133L	S322G								
3	+++	2624/2654	C73R	A83V	T135I								
4	++	2624/2654	C73R	E177G									
5	++	2624/2654	C73R										
6	+/-	2624/2654	C153Y	E197K	T281A								
7	+	2624/2654	C73R	T256A	N466S								
8	+++	2624/2654	C73R	E141V									
9	++	2624/2654	C73R	E303K									
10	+++	2624/2654	Q41K										
16	+++	2680/26886	Q41K	P16A	G23S	T9M	Q362E						
19	+/-	2700/2701	R21H	S34A	Q80H	A84S	E303D	H374D	A440T	A441V	C445S	P499S	
20	+	2700/2701	F31L	T409I									
21	+++	2700/2701	F31L	M97I	E113D	D146N	P169S	N367I	H458Y				
30	+/-	2681/26886	I43V	Q295L									
31	++	2680/26886	F31L	P101S	C153R	C159S	E162K	R293L	S311N				
32	++	2680/26886	L14I	E18V	G138C	E338G	V361L	P389S	N400S				
33	++	2680/26886	Q41R	S174Y	A402T								
34	++	2680/26886	F31L	T52I	P108S	V111I							
36	+/-	2700/2701	D85N	I143F	M232I	T315I	S382Y	M385K					
37	++	2700/2701	T46I	Q62R	K77R	S92C	N367Y	V373I					
38	-/+	2700/2701	Q41R	T294I	P310L	G337D	P389L	A394V	G436S				
39	+	2680/26886	T52N	V111I	T139	V184I	T281A						
40	+++	2680/26886	Q41R	D4E	V87I	D110E	E141K	A189T	N276D	T347R	N367I	Q377R	A425T
41	-/+	2680/26886	D131N	S133L	R312G	A429G							
wild-type	-		N/A	N/A									

5

Table 4 summarizes amino acid substitutions that were isolated multiple times, suggesting that they are particularly important for improving *lovE* variant activity on *lovFp-HIS3p-neo* expression.

10

**Table 4: *lovE* Mutations Isolated Multiple Times**

Amino Acid Change	Number of Times Isolated in <i>lovE</i> 1-41	<i>lovE</i> variant
F31L	4	20, 21, 31, 34
Q41K	2*	10, 16
Q41R	3*	33, 38, 40
T52I/T52N	1 each	34, 39
C73R	6*	3, 4, 5, 7, 8, 9
P101S/P101Q	1 each	31, 34
V111I	2	34, 39
S133L	2	2, 41
E141V, E141K	1 each	8, 40
C153Y/C153R	1 each	6, 31
T281A	2	6, 39
N367I/N367Y	2/1	21, 40, 37
P389S/P389L	1 each	32, 38

\* allele was isolated in additional *lovE* variants that were not fully sequenced

**Example 5: Increased *lovF-lacZ* Expression in *S. cerevisiae***

15

In order to quantify the increase in *lovF* expression,  $\beta$ -galactosidase activity was measured in *lovE* variant transformed *S. cerevisiae* strains that also harbored *lovFp-lacZ* reporter derivative plasmids. *lovF-lacZ* reporter derivative plasmids were constructed as follows.

20

Plasmid MB1918 contains the *lovFp-lacZ* reporter gene. It can be derived from pRS424 (Sikorski and Hieter (1989) *Genetics* 122:19-27). First, primers MO107 (SEQ ID NO:25) and MO197 (SEQ ID NO:26) are used to PCR amplify the *lacZ* gene from Yep355 (Myers, et al., *Gene*

5   **45**:299-310 (1986)). This lacZ-containing fragment was  
 inserted into the *Bam*HI-*Hind*III sites of pRS416 (Sikorski  
 and Hieter, *Genetics* **122**:19-27 (1989)). This same lacZ  
 fragment can be cut out of the resulting vector with  
*Kpn*I-*Not*I and inserted into the same sites of pRS424 to  
 10 create pRS424-lacZ. Primers M01293 (SEQ ID NO:27) and  
 M01294 (SEQ ID NO:28) are used to PCR amplify a 2.09 kb  
 fragment of the *lovF* promoter from *A. terreus* genomic  
 DNA. The *lovF* promoter fragment was then cut with *Not*I-  
*Bgl*II and inserted into *Not*I-*Bam*HI linearized pRS424-  
 15 lacZ.

Plasmid MB2114 contains the *lovFp*-*CYC1p*-lacZ reporter gene. It can be derived from pRS424-lacZ (see MB1918 plasmid construction). Primers M01787 (SEQ ID NO:29) and M01788 (SEQ ID NO:30) are used to amplify the 20 264 bp basal *CYC1* element from pRS415 *CYC1* (Mumberg, et al., *Gene* **156**:119-122 (1995)). This 264 bp fragment was inserted upstream of the pRS424-lacZ derivative which has been digested with *Spe*I-*Bam*HI. Finally, the *lovF* promoter from MB1918 was PCR amplified with M01793 (SEQ 25 ID NO:31) and M01794 (SEQ ID NO:32) and inserted into the *Not*I-*Spe*I sites to create MB2114.

Yeast strains utilized in this study include strains MY2145 and MY2159, which are both derived from the *S. cerevisiae* sigma 1278b strain background; the genotypes 30 are both strains are as follows: *MATA ura3Δ0 leu2Δ0 his3Δ::hisG trp1Δ0::hisG*. MY2145 and MY2159 contain the *lovFp*-lacZ reporter plasmids MB2114 and MB1918, respectively.

MY2124 transformed with individual *lovE* variant 35 plasmids was mated to *S. cerevisiae* strains MY2154 and MY2159. Diploids were selected on SC-UraTrp media. Multiple diploids from each individual mating were assayed for *lovFp*-lacZ expression using 96 well format  $\beta$ -

5 galactosidase assays. For  $\beta$ -galactosidase assays, cells were transferred from transformation plates to 96-well microtiter plates containing 200  $\mu$ l Z buffer. 12 strains were transferred simultaneously using a 12-channel multi-pipettor to scoop cells from transformation plates.

10 Duplicate samples were prepared for all assays. OD<sub>600</sub> readings were taken on samples in Z buffer. These values were used to normalize for equal cell number in all assays. After determining OD<sub>600</sub>, 150  $\mu$ l of each sample in Z buffer was transferred onto a Millipore Multiscreen

15 Assay System (Nitrocellulose Immobilon NC), filtered, and then washed by filtering 200  $\mu$ l Z buffer. 100  $\mu$ l Z buffer with  $\beta$ ME and detergents was then added to each well, as was 20  $\mu$ l 4 mg/ml ONPG. Reactions were incubated at 30°C, stopped with 50 $\mu$ l 1 M Na<sub>2</sub>CO<sub>3</sub>, filtered

20 into a polystyrene 96-well assay plate, and OD<sub>420</sub> was determined for each assay well.  $\beta$ -galactosidase units were determined using the Miller formula (O.D. 420 X 1000) / (OD<sub>600</sub>\*minutes\*volume in mL). Z buffer is made by dissolving the following in 1 L of water (16.1 g Na<sub>2</sub>HPO<sub>4</sub>-7H<sub>2</sub>O, 5.5g NaH<sub>2</sub>PO<sub>4</sub>-H<sub>2</sub>O, 0.75 g KCl and 0.246 g MgSO<sub>4</sub>-7H<sub>2</sub>O).

25 Z buffer with detergents and  $\beta$ ME is made as follows: 9.8 ml Z buffer, 100  $\mu$ l 20 mg/ml CTAB, 100  $\mu$ l 10 mg/ml sodium deoxycholate, and 69  $\mu$ l  $\beta$ ME. Control plasmids utilized in these studies included MB968, MB2478 and MB1644.

30 Results of these studies are presented in Figures 2-5, demonstrating increased transcription-activating properties of the *lovE* variants disclosed herein.

**Example 6: Secondary Metabolite Production**

5 Transformation of filamentous fungi was performed according to the following procedure. Protoplasts were generated by inoculating rich media with spores. Spores were allowed to germinate for about 20 hrs or until germ tubes were between 5 and 10 spore lengths. The germlings  
10 were centrifuged and washed twice with sterile distilled water and once with 1 M magnesium sulfate. Germlings were then resuspended in 1M magnesium sulfate containing approximately 2 mg/ml of Novozyme. Tubes were then incubated at 30°C shaking at 80 RPM for about 2 hrs or  
15 until most of the hyphae were digested and protoplasts were abundant. Protoplasts were filtered through one layer of Miracloth. At least one volume of STC was added and protoplasts were centrifuged. Protoplasts were washed twice with STC. Protoplasts then were resuspended  
20 in 1ml STC and counted in a hemacytometer. A final concentration of approximately  $5 \times 10^7$  protoplasts/ml were frozen in a 9:1:0.1 solution of STC, SPTC and DMSO in a Nalgene Cryo cooler at -80°C (cools -1°C/min).

Solutions for transformation were as follows: STC  
25 (0.8 M Sorbitol, 25 mM Tris-HCl pH 7.5, 25 mM CaCl<sub>2</sub>) and SPTC (0.8 M Sorbitol, 40% PEG 4000, 25 mM Tris-HCl pH 8, 50 mM CaCl<sub>2</sub>). Transformation was accomplished according to the following protocol. 1-5 µg of DNA comprising a *lovE* variant according to the invention in a fungal  
30 expression vector was placed in a 50 ml Falcon tube. 100 µl of previously frozen protoplasts were added to the DNA, gently mixed, and then incubated on ice for 30 min. 15 µl of SPTC was added, followed by mixing by tapping and incubation at RT for 15 min. 500 µl SPTC was added  
35 and mixed well by tapping and rolling, then incubated at RT for 15 min. 25 mls of regeneration minimal medium was added, mixed well and poured on plates containing 25 mls

5 of regeneration minimal medium with 2X the concentration of selection drug.

Transformation plates were incubated at 26°C for 5-6 days or until colonies started to appear. Regeneration minimal medium contains trace elements, salts, 25 mM

10 sodium nitrate, 0.8 M Sucrose, and 1% agarose at pH 6.5.

The selection drug that was used successfully with *A. terreus* is phleomycin, a broad-spectrum glycopeptide antibiotic. Transformants were picked onto new plates with a toothpick (if the fungus was sporulating) or with

15 sterile forceps (if the fungus did not sporulate).

Purification plates contained minimal medium (same as regeneration minimal medium but containing 2 % instead of 0.8 M sucrose) and 1X drug concentration. Picked transformants were incubated at 26°C for 5-6 days.

20 Transformants were grown in production media for secondary metabolite production. Briefly, for *A. terreus* and lovastatin production, spores were used as the

inoculum. Spores were obtained from the purification plate by using a wooden inoculation stick. The medium

25 was RPM containing corn steep liquor, sodium nitrate, potassium phosphate, magnesium sulfate, sodium chloride, P2000 (Dow chemical), trace elements and lactose or glucose as carbon source. The medium was pH 6.5. Flasks were incubated at 26°C with shaking at 225 RPM.

30 For static 96-well cultures, the same medium was used and the spores were obtained from the purification plate with a wooden toothpick. 96-well plates were incubated, without shaking at 26°C.

Sampling was done after after 5 days for 35 lovastatin. For shake flask experiments 1-1.5 mls of supernatant was placed into 96-well plates, which were centrifuged and supernatants transferred to new 96-well plates. Samples were frozen at -80°C for storage or for later assays.

5        Cultures that were grown standing in a 96-well plate  
were centrifuged and the supernatant was transferred to a  
new 96 well plate. Samples were frozen at -80°C.

**Example 7: Measurement of Secondary Metabolite Production**

10       The concentration of the secondary metabolite  
lovastatin was determined by enzyme inhibition assay  
(Figure 6). Briefly, 10  $\mu$ L of sample was removed and  
diluted 1:100 in H<sub>2</sub>O. 10  $\mu$ l of this diluted broth was  
assayed in a reaction (200  $\mu$ L total) containing 1 mM  
15 HMGCoA, 1 mM NADPH, 0.005 mM DTT and 5  $\mu$ l (His)<sub>6</sub>HMGR. The  
disappearance of absorbance at 340 nm was observed over  
time. This represents the disappearance of NADPH, and  
lovastatin inhibits this reaction.

20       The initial velocities were calculated for the  
reactions containing samples, adjusted for dilution, and  
compared to reactions containing lovastatin standards to  
determine levels of metabolite produced. (His)<sub>6</sub>HMGR was  
expressed in *Saccharomyces cerevisiae* and purified with a  
nickel column.

25       The results from ten individual transformants for  
each allele are shown in standard box plot format in  
Figure 6. Lovastatin concentration from the  
corresponding wild-type *lovE* control is shown in matching  
fill pattern. For example, *lovE* alleles 2, 7, 8 and 9  
30 were all transformed and assayed at the same time as the  
non-hatched wild-type control. The horizontal line in  
each individual box represents the median.

35       Lovastatin concentration was also determined by high  
pressure liquid chromatography (HPLC). Briefly, 100  $\mu$ L  
of broth sample was removed and diluted 1:10 into 70% H<sub>2</sub>O-  
30% acetonitrile (900  $\mu$ l). This mixture was spun down to  
pellet debris at 13000 RPM for 5 minutes. 900  $\mu$ l of this

5 diluted broth was transferred to a vial and the sample  
 was analyzed by HPLC. 10  $\mu$ l were injected into a Waters  
 HPLC system (996 photo-diode array detector, 600 E pump  
 controller and 717 autosampler) equipped with a YMC-Pack  
 ODS column (Aq-302-3, 150 x 4.6 mm ID, 5-3  $\mu$ M pore size)  
 10 and eluted with isocratic 40% aqueous acetic acid (0.7%) -  
 60% acetonitrile for 8 minutes. Lovastatin was detected  
 at 238 nm to have a retention time of 6.5 minutes and was  
 quantified using a calibration curve created from pure  
 lovastatin samples.

15 The results from ten individual transformants for  
 each *lovE* variant are shown in standard box plot format  
 in Figure 7A and 7B. Thirty individual wild-type *lovE*  
 transformants and ten individual MB2143 negative control  
 transformants were tested. Identical controls are  
 20 plotted in Figures 7A and 7B.

PCR analysis of *A. terreus* transformants  
 demonstrates that greater than fifty percent of the  
 transformants contain the transgene. Variability in  
 levels of transgene expression can presumably be  
 25 influenced by integration site and copy number. *lovE*  
 variants containing identical amino acid substitutions  
 are labeled.

The amino acid and nucleic acid sequences of *lovE*  
 variant sequences are presented in Table 5 and Table 6,  
 30 respectively.

Table 5: Amino Acid Sequences of Variants of the <i>lovE</i> Gene	
<i>lovE-1</i>	maadqgifsntsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsqrflppdsy dwshtsigtdeaidtdcwglsqcdggfscqleptlpd1pspfestvekaplppvssdiaraasaq relfddlsavsqeleeillavtvevpkqeiwthpigmffnasrrlltv1rqqaqadcrqgt1dec 25 lrtknlfavhcyilnvriltaisellisqirrtqnshmsplegsrsqspqrddtssssghssvd tipffsenlpigelfpyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg ediartgatnsarceeqpttpaarvlfmflsdegafeakeaksagsrgrtiaalrrcyedifslark hkhgmlrdlnnipp (SEQ ID NO:41)

**love-2**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsrfldppdsy  
 dwlwtsgtdeaidtgcwglscdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkgeiwhpigmffnasrrlltvrlrqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvqllreneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:42)

**love-3**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsrfldppdsy  
 dwswtsgtdeaidtgcwglscdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkgeiwhpigmffnasrrlltvrlrqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvqllreneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:43)

**love-4**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsrfldppdsy  
 dwswtsgtdeaidtgcwglscdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkgeiwhpigmffnasrrlltvrlrqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvqllreneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:44)

**love-5**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsrfldppdsy  
 dwswtsgtdeaidtgcwglscdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkgeiwhpigmffnasrrlltvrlrqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvqllreneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:45)

**love-6**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsrfldppdsy  
 dwswtsgtdeaidtgcwglscdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 rklfddlsavsqeleeillavtvewpkgeiwhpigmffnasrrlltvrlrqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvqllreneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:46)

**love-7**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqrsaadlvsadpdpc1hmssppvpsqlpldvseshssntsrfldppdsy  
 dwswtsgtdeaidtgcwglscdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkgeiwhpigmffnasrrlltvrlrqqaqadchqgaldec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvqllreneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnsipp (SEQ ID NO:47)

**LOVE-8**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqsraadlvsadpdpc1hmssppvpsqlpldvseshssntsrfqfldppdsy  
 dswstsigtdeaidtdcwglscqcdggfscqleptlpdlspsfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgaldec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnsipp (SEQ ID NO:48)

**LOVE-9**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqsraadlvsadpdpc1hmssppvpsqlpldvseshssntsrfqfldppdsy  
 dswstsigtdeaidtdcwglscqcdggfscqleptlpdlspsfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgaldec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnsipp (SEQ ID NO:49)

**LOVE-10**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvyserrpkrklrqsraadlvsadpdpc1hmssppvpsqlpldvseshssntsrfqfldppdsy  
 dswstsigtdeaidtdcwglscqcdggfscqleptlpdlspsfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgaldec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnsipp (SEQ ID NO:50)

**LOVE-11**

maadqgiftmnsvtlsavegsrtsgtlprrafrrscdrchakkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqsraadlvsadpdpc1hmssppvpsqlpldvseshssntsrfqfldppdsy  
 dswstsigtdeaidtdcwglscqcdggfscqleptlpdlspsfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:51)

**LOVE-12**

maadqgiftnsvtlspvegshtggtlprrafrracdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrhsrasdlvsadpdpc1hmssppvpsqlpldvseshssntsrfqfldppdsy  
 dswstsigtdeaidtdcwglscqcdggfscqleptlpdlspsfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmspldgsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvdsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtitvlrrsyedifslark  
 hkhgmlrdlnnips (SEQ ID NO:52)

**LOVE-20**

maadqgiftnsvtlspvegsrtggtlprralrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqsraadlvsadpdpc1hmssppvpsqlpldvseshssntsrfqfldppdsy  
 dswstsigtdeaidtdcwglscqcdggfscqleptlpdlspsfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspsrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpitpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:53)

**love-21**

maadqgiftnsvtlspvegsrtggtlprralrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hissppvpsqslpldvsseshssntsrqfldppdsy  
 dwsmtsigtdeaidtncwglsqcdggfscqlestlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reieitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkgmlrdlnnipp (SEQ ID NO:54)

**love-30**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkvkctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqslpldvsseshssntsrqfldppdsy  
 dwsmtsigtdeaidtdcwglsqcdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtlnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkgmlrdlnnipp (SEQ ID NO:55)

**love-31**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqslpldvsseshssntsrqfldppdsy  
 dwsmtsigtdeaidtdcwglsqrdggfssqlkptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkgmlrdlnnipp (SEQ ID NO:56)

**love-32**

maadqgiftnsvtispvvgsgrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqslpldvsseshssntsrqfldppdsy  
 dwsmtsictdeaidtdcwglsqcdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigglfsyvdplthalfsacttlhvgql1reneitlgvhsaqgiaasismsgesg  
 ediartgatssarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkgmlrdlnnipp (SEQ ID NO:57)

**love-33**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrcharkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmssppvpsqslpldvsseshssntsrqfldppdsy  
 dwsmtsigtdeaidtdcwglsqcdggfscqleptlpd1pspfeytvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnstrceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkgmlrdlnnipp (SEQ ID NO:58)

**love-34**

maadqgiftnsvtlspvegsrtggtlprralrrscdrchaqkikctgnkevigrapcqrcqqagl  
 rcvysercpkrklrqrsaadlvsadpdpc1hmsspqvpsqslsldiseshssntsrqfldppdsy  
 dwsmtsigtdeaidtdcwglsqcdggfscqleptlpd1pspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtvewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaiselllsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaqgiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkgmlrdlnnipp (SEQ ID NO:59)

**love-36**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqsraanlvsadpdpc1hmssppvpsqslpldvseshssntsqrqfldppdsy  
 dswtsigtdeafdtcwglsgcdggfscqleptlpdlpspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtewpkqeiwthpigiffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspqrddisssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaggiaayisksgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:60)

**love-37**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikcignkevtgrapcqrcqragl  
 rcvysercpkrrlrqsraadlvsadpdpc1hmssppvpsqslpldvseshssntsqrqfldppdsy  
 dswtsigtdeaidtgcwglsgcdggfscqleptlpdlpspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspqrddtssssghscvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reyeitlgvhsaggiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:61)

**love-38**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrcharkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqsraadlvsadpdpc1hmssppvpsqslpldvseshssntsqrqfldppdsy  
 dswtsigtdeaidtgcwglsgcdggfscqleptlpdlpspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirriqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpidelfsyvdplthalfsacttlhvgvql1reneitlgvhsaggiaasismsgelg  
 edivrtgatnsarceeqpttpaarvlfmflsdeafqeqsagsrsrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:62)

**love-39**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrcharkikctgnkevngrapcqrcqqagl  
 rcvysercpkrklrqsraadlvsadpdpc1hmssppvpsqslpldiseshssntsqrqfldppdsy  
 dswtsigideaidtgcwglsgcdggfscqleptlpdlpspfestvekaplppissdiaraasaq  
 relfddlsavsqeleeillavtewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirriqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaggiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:63)

**love-40**

maaeqgiftnsvtlspvegsrtggtlprrafrrscdrcharkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqsraadlisadpdpc1hmssppvpsqslplevseshssntsqrqfldppdsy  
 dswtsigtdkaidtgcwglsgcdggfscqleptlpdlpspfestvekaplppvssditraasaq  
 relfddlsavsqeleeillavtewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyildvrltaisellsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplrhalfsacttlhvgvql1reieitlgvhsargiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdegtfpeaksagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:64)

**love-41**

maadqgiftnsvtlspvegsrtggtlprrafrrscdrchaqkikctgnkevtgrapcqrcqqagl  
 rcvysercpkrklrqsraadlvsadpdpc1hmssppvpsqslpldvseshssntsqrqfldppdsy  
 nwltwtsigdeaidtgcwglsgcdggfscqleptlpdlpspfestvekaplppvssdiaraasaq  
 relfddlsavsqeleeillavtewpkqeiwthpigmffnasrlltv1rqqaqadchqgt1dec  
 lrtknlftavhcyilnvriltaisellsqirrtqnshmsplegsrsqspqrddtssssghssvd  
 tipffsenlpigelfsyvdplthalfsacttlhvgvql1reneitlgvhsaggiaasismsgepg  
 ediartgatnsarceeqpttpaarvlfmflsdeafqeqsagsrgrtiaalrrcyedifslark  
 hkhgmlrdlnnipp (SEQ ID NO:65)

Table 6: DNA Sequences of Variants of the *lovE* Gene***lovE-1***

ATGGCTGCAGATCAAGGTATATTACGAACCTGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCCGCCGTGCATTCCGACGCTTGTGATCGGTGTATGCACAAAGATCA  
 AATGTAATGGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCACTGCCCCACAGAGCTTGCAC  
 TAGACGTATCCGAGTCGATTCTCAAATACCTCCGGCAGTTCTTGTATCCACCGGACAGCTAC  
 GACTGGCTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGCTCCA  
 ATGTGATGGAGGCTTCAGCTGTCAAGTTAGGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGCAC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCAGGAACCTGGAAAGAGATCCTCTGGCCGTGAC  
 GGTAGAAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTACGAC  
 GGCTTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCGTCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCAAAATTAGGCGGACCCAGAACAGCCATATGAGCCCAGTGGAAAG  
 GGAGTCGATCCAGTCGCCAGAGACAGCACCCAGCAGCAGCGGGCCACAGCAGTGTGAC  
 ACCATACCCTCTTACCGAGAACCTCCCTATTGGTGGAGCTGTTCCATGTTGACCCCTGAC  
 ACACGCCCTATTCTCGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTGCTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGGCCACACTCCAGC  
 GGCTCGGGTTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCCATGA (SEQ ID NO: 66)

***lovE-2***

ATGGCTGCAGATCAAGGTATATTACGAACCTGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCCGCCGTGCATTCCGACGCTTGTGATCGGTGTATGCACAAAGATCA  
 AATGTAATGGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCACTGCCCCACAGAGCTTGCAC  
 TAGACGTATCCGAGTCGATTCTCAAATACCTCCGCAATTGCTGAGTCCACCGGACAGCTAC  
 GACTGGTTGAGGCTTCAGCTGTCAAGTTAGGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 ATGTGATGGAGGCTTCAGCTGTCAAGTTAGGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGCAC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCAGGAACCTGGAAAGAGATCCTCTGGCCGTGAC  
 GGTAGAGTGGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTACGAC  
 GGCTTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCAAAATTAGGCGGACCCAGAACAGCCATATGAGCCCAGTGGAAAG  
 GGAGTCGATCCAGTCGCCAGAGACAGCACCCAGCAGCAGCGGGCCACGGCAGTGTGAC  
 ACCATACCCTCTTACCGAGAACCTCCCTATTGGTGGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTCGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTGCTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGGCCACACTCCAGC  
 GGCTCGGGTTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCCATGA (SEQ ID NO: 67)

**love-3**

ATGGCTCAGATCAAGGTATATTACGAACCTGGTCACTCTCGCCAGTGGAGGGTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGCCTCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGACGCCAAGCGCAAGCTACGCCAATCCAGGGTAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCCGCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGCGTGGATCTGACGGACTGACGAGGTTATTGACACTGACTGCTGGGGCTGTCCTCA  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCGCCAGTGC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAGAGATCCTCTGCCGTGAC  
 GGTAGAATGCCGAAGCAGGAAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCGCCAAACAGCGCAGGCCACTGCCATCAAGGCACACTAGAGGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGGCCACTGGAA  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGGCAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGATGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGCAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:68)

**love-4**

ATGGCTCAGATCAAGGTATATTACGAACCTGGTCACTCTCGCCAGTGGAGGGTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGCCTCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGACGCCAAGCGCAAGCTACGCCAATCCAGGGAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCCGCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGTTATTGACACTGACTGCTGGGGCTGTCCTCA  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCGCCAGTGC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAGAGATCCTCTGCCGTGAC  
 GGTAGAATGCCGAAGCAGGAAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCGCCAAACAGCGCAGGCCACTGCCATCAAGGCACACTAGAGGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGGCCACTGGAA  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGGCAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGATGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGCAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:69)

**love-5**

ATGGCTGCAGATCAAGGTATATTACGAACCTGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGCCTCGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGACGCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCCGCCAGTGCCTCACAGAGCTTGCCGC  
 TAGACGTATCCGAGTCGATTCCCAAATACCTCCCGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCA  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTAAAAAGCTCCGTTGCCACCCGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGAC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCGCCAAACAAGCGCAGGCCAGTGCCTCACAGGACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACGACACCAGCAGCAGCAGCGGCCAGCAGTGTGAC  
 ACCATACCCCTTCTTACGAGAACCTCCCTATTGGTGAGCTGTTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCCATCAGCATGAGGGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACTACTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGCAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATTCCCTCCATGA (SEQ ID NO:70)

**love-6**

ATGGCTGCAGATCAAGGTATATTACGAACCTGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGCCTCGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCCGCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGATTCCCAAATACCTCCCGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCA  
 ATATGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTAAAAAGCTCCGTTGCCACCCGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAAAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGAC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCGCCAAACAAGCGCAGGCCAGTGCCTCACAGGACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGGCCGC  
 CATATCGGAGTTGCTCCTGCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGAGAACCTCCCTATTGGTGAGCTGTTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCCATCAGCATGAGGGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACCTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGCAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATTCCCTCCATGA (SEQ ID NO:71)

**love-7**

ATGGCTCAGATCAAGGTATATTACGAACCTCGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGTCCCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGACGCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCCGCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGTCGGACCTCGATTGGCACTGACGAGGTTATTGACACTGACTGCTGGGGCTGTCCA  
 ATGTGATGGAGGCTTCAGCTGTAGTTAGGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGC  
 GGTAGAATGCCGAAGCAGGAAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCGCCAAACAGCGCAGGCCACTGCCATCAAGCGCAGTACAGGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAA  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTGCTTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTTGAGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCGCAAA  
 CACAAACATGGCATGCTCAGAGACCTAACAGTATTCCCTCCATGA (SEQ ID NO:72)

**love-8**

ATGGCTCAGATCAAGGTATATTACGAACCTCGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGTCCCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGACGCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCCGCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGTCGGACCTCGATTGGCACTGACGTGGTATTGACACTGACTGCTGGGGCTGTCCA  
 ATGTGATGGAGGCTTCAGCTGTAGTTAGGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGC  
 GGTAGAATGCCGAAGCAGGAAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCGCCAAACAGCGCAGGCCACTGCCATCAAGGCACACTAGAGGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAA  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTGCTTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTTGAGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCGCAAA  
 CACAAACATGGCATGCTCAGAGACCTAACAAATTCCCTCCATGA (SEQ ID NO:73)

**love-9**

ATGGCTCAGATCAAGGTATATTACGAACCTCGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGTCCCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGACGCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 TTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCTCCAGTGCCCTCACAGAGCTGCCAC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCCA  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAGCTCCGTTGCCACCCGATCGAGCGACATTGCTCGTGCGCCAGTGCAGCAA  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCCGTGAC  
 GGTAGAATGCCGAAGCAGGAAATCTGGACCCATCCCCTCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCCAACAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAACGTGCCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGAAAG  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTGAGCTTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGATGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCCATGA (SEQ ID NO:74)

**love-10**

ATGGCTCAGATCAAGGTATATTACGAACCTCGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGTCCCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCTCCAGTGCCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCCA  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAGCTCCGTTGCCACCCGATCGAGCGACATTGCTCGTGCGCCAGTGCAGCAA  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCCGTGAC  
 GGTAGAATGCCGAAGCAGGAAATCTGGACCCATCCCCTCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCCAACAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTCGGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGCTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTGAGCTTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGATGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCCATGA (SEQ ID NO:75)

**love-16**

ATGGCTGCAGATCAAGGTATATTCACTGAACACTCGGTCACTCTCTGCAGTGGAGGGTTACGCAC  
 CAGTGGAACATTACCCGCCGTGCATTCCGACGCTCTGTGATCGGTGTCATGCACAAAAGATCA  
 AATGTAATGGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTGAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCTCGCCTCCAGTGCCTCACAGAGCTGCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTC  
 ATGTAATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACAGTTGAAAAAGCTCCGTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGAGGAACTGGAAGAGAGATCCTCTGGCGTGC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCATCGGAATGTTTTCAATGCGTCAGC  
 GGCTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTATCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACACTAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTAGAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGGAGTACACTCCGCCAGGGCATTGCACTCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACACTCCAGC  
 GGCTGGGTTTTGTCATGTTCTTGAGTGATGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO: 76)

**love-19**

ATGGCTGCAGATCAAGGTATATTCACTGAACACTCGGTCACTCTCGCCAGTGGAGGGTTACACAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCGCTCTGTGATCGGTGTCATGCACAAAAGATCA  
 AATGTAATGGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTGAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCATTCCAGGGATCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCTCGCCTCCAGTGCCTCACAGAGCTGCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTC  
 ATGTAATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGAGGAACTGGAAGAGAGATCCTCTGGCGTGC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCATCGGAATGTTTTCAATGCGTCAGC  
 GGCTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACACCCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTATTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGGAGTAGACTCCGCCAGGGCATTGCACTCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTGGGTTTGTCATGTTCTTGAGTGATGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCACAGTACTGCGACGAGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO: 77)

**love-20**

ATGGCTGCAGATCAAGGTATATTACGAACCTCGGTCACTCTCTGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCACTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTACTGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCTCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGCCCC  
 ATGTGATGGAGGCTTCAGCTGCAAGTTAGAGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGCGCAGGAACCTGGAAGAGAGATCCTCTGGCGTGC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCCCTCGGAATTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCTGCCAACAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCCACCTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACACCCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTCGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCAGCATGAGCGGGAAACCGAGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGGCCACTCCAGC  
 GGCTGGGTTTTGTCATGTTCTGAGTGATGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO: 78)

**love-21**

ATGGCTGCAGATCAAGGTATATTACGAACCTCGGTCACTCTCTGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCACTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTACTGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATATCCTCGCCTCAGTGCCTCACAGAGCTTACCGC  
 TAGACGTATCCGATTGCACTCCTCAAATACCTCCGGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTAATGCTGGGGCTGCCCC  
 ATGTGATGGAGGCTTCAGCTGCAAGTTAGAGTCACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGCGCAGGAACCTGGAAGAGAGATCCTCTGGCGTGC  
 GGTAGAATGGCTAAGCAGGAAATCTGGACCCATCCCCTCGGAATTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCTGCCAACAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCCACCTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACACCCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTCGGCTTGCACTACGTTACATGTTGGGTACAATTGCTCGTGAGATTGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCAGCTTCCATCAGCATGAGCGGGAAACCGC  
 GAGGATATAGCCAGGACAGGGCAACCAATTCCGCAAGATGCGAGGAGCAGCCACCTCCAGC  
 GGCTGGGTTTTGTCATGTTCTGAGTGATGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGAAA  
 CACAAATATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO: 79)

**LOVE-30**

ATGGCTCAGATCAAGGTATATTACGAACCTCGTCACTCTCTGCCAGTGGAGGGTCACGCAC  
 CGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGCCTCTGTCAAGCGTGCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCGCCTCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGATTCCCAAATACCTCCGGCAATTCTTGTGATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCA  
 ATGTGATGGAGGTTAGCTGAGTGTAGAGCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTACGAC  
 GGCTTCTTACTGTCCTGCGCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGCGGACCCCTGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTCGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAAGATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGATTGAGCTTCCAGGAGGCAAGTCTGTTGAGGAAACCAGGC  
 GAGGATATAGCCAGGAACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGCAAGTCTGTTGAGGAAACCAGGC  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCCGAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:80)

**LOVE-31**

ATGGCTCAGATCAAGGTATATTACGAACCTCGTCACTCTCTGCCAGTGGAGGGTCACGCAC  
 CGTGGAACATTACCCGCCGTGCATTACGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCGTGCCTCTGTCAAGCGTGCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGTTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCGCCTTCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGATTCCCAAATACCTCCGGCAATTCTTGTGATCCACCGGACAGCTAC  
 GACTGGCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCA  
 ACGTGTGGAGGTTAGCTCTCAGTTAAAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTACGAC  
 GGCTTCTTACTGTCCTGCGCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTACTGTCGCAAATTAGGCTGCCAGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCGAACAGAGACGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTCGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAAGATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGATTGAGCTTCCAGGAGGCAAGTCTGTTGAGGAAACCAGGC  
 GAGGATATAGCCAGGAACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGCAAGTCTGTTGAGGAAACCAGGC  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCCGAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:81)

**love-32**

ATGGCTGCAGATCAAGGTATATTCACTAAGTCGGTCACTATCTGCCAGTGGGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAAGTGGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCTGCCAGTGCCTCACAGAGTTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGTCGGACCTCGATTTGACTGACGAGGTTATTGACACTGACTGCTGGGGCTGTCCA  
 ATGTGATGGAGGTTAGCTGAGTGTAGAGCCAAGCGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTACGAC  
 GGCTTCTTACTGTCCCTGCGCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGCGGACCCAGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACCGAGAACCTCCCTATTGGTGGGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGCTACAATTGCTCGTGAAGATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCCATCAGCATGAGCGGGAAATCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAAGTCCGCAAGATGCGAGGAGCAGCCGACACTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:82)

5

**love-33**

ATGGCTGCAGATCAAGGTATATTCACTAAGTCGGTCACTCTCGCCAGTGGGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAAGTGGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCTGCCAGTGCCTCACAGAGCTTGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGTATCCACCGGACAGCTAC  
 GACTGGTCGGACCTCGATTGGCACTGACGAGGTTATTGACACTGACTGCTGGGGCTGTCCA  
 ATGTGATGGAGGTTAGCTGAGTGTAGAGCCAAGCGCTGCCGGATCTACCTTCGCCCTCGAGT  
 ATACGGTTAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTACGAC  
 GGCTTCTTACTGTCCCTGCGCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCAGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGCGGACCCAGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCGAGCAGAGACGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTACCGAGAACCTCCCTATTGGTGGGCTACAATTGCTCGTGAAGATGAGA  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGCTACAATTGCTCGTGAAGATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCCATCAGCATGAGCGGGAAACCGAC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCACAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGTGAAGGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:83)

**love-34**

ATGGCTCAGATCAAGGTATATTACGAACCTCGGTCACTCTCGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTGCGACGCTTGTATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTATTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTATACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCTCAAGTGCCTCACAGAGCTGTCGC  
 TAGACATATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCTCA  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTAAAAAGCTCCGTTGCCACCCGTATCGAGCGACATTGCTCGTGCGCCAGTGCGBAA  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGCCGTGAC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCCTCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCCAACAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTCGGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCAAATTAGGCGGACCCAGAACAGCCATATGAGGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTCTTACCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCAGCTTCCATCAGCATGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACACTCCAGC  
 GGCTCGGGTTTGTTCATGTTCTGAGTGATGAAGGGCATTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCTGCCCGCAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:84)

**love-36**

ATGGCTCAGATCAAGGTATATTACGAACCTCGGTCACTCTCACCAAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTATCGGTGTCATGCACAAAGATCA  
 AATGTAATGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGAATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTTACACATGTCCCGCCTCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCTCAAATACCTCCGGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGGACCTCGATTGGCACTGACGAGGCTTTGACACTGACTGCTGGGGCTATCCCA  
 ATGTGATGGAGGCTTCAGCTGTCAGCTAGAGCCAACGCTGCCGGATCTACCTTCGCCCTCGAGT  
 CTACGGTTAAAAAGCTCCGTTGCCACCCGTATCGAGCGACATTGCTCGTGCGCCAGTGCGBAA  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGCCGTGAC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCCTCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCCAGCAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTCGGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGCAAATTAGGCGGACCCAGAACAGCCATATGAGGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACGACATCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTCTTACCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCAGCTTACATCAGCAAGAGCGGGAAACCAGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACACTCCAGC  
 GGCTCGGGTTGTTCATGTTCTGAGTGATGAAGGGCATTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCTGCCCGCAAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:85)

**love-37**

ATGGCTGCAGATCAAGGTATATTACGAACCTCGTCACTCTCTGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTATTGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAGCGTTGCCAACGGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAGGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGTTGCACATGTCCTCGCCCTCAGTGCCCTCACAGAGCTTGC  
 TAGACGTATCCGAGTCGATTCTCAAATACTCCCGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCC  
 ATGTGATGGAGGCTTCAGCTGCACTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGC  
 CGAGAGCTTTGATGACCTGCGGGTGTGCAGGAACGGAACTGGAAGAGATC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCCAGTGGAAATGTT  
 GGCTTCTTACTGTCCTGCGCCAACAAGCTCAGGCCACTGCCATCAAGGC  
 ACAGCTACAGCTGAGGAGACACCAGCAGCAGCAGCGGCCACAGCTGTG  
 ACCATACCCCTTTAGCGAGAACCTCCCTATTGCTGAGCTTCTCTATGTT  
 ACACAGCCCTATTCTGGCTTGACTACGTTACATGTTGGGTACAATTGCT  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGC  
 GGCTGGGTTTTGTCATGTTCTGAGTGAAGGGCTTCCAGGAGGCAAAGT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATAT  
 CACAAACATGGCATGCTCAGAGATCTCAACAATATTCCCATGA (SEQ ID NO: 86)

**love-38**

ATGGCTGCAGATCAAGGTATATTACGAACCTCGTCACTCTCTGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTACTGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAGCGTTGCCAGCAAGCTGGACTT  
 CGATGCGTCTATAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGTTGCACATGTCCTCGCCCTCAGTGCCCTCACAGAGCTTGC  
 TAGACGTATCCGAGTCGATTCTCAAATACTCCCGCAATTCTTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACGAGGCTATTGACACTGACTGCTGGGGCTGTCCC  
 ATGTGATGGAGGCTTCAGCTGCACTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAAGCTCCGTTGCCACCGGTATCGAGCGACATTGCTCGTGC  
 CGAGAGCTTTGATGACCTGCGGGTGTGCAGGAACGGAAAGAGATC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCCAGTGGAAATGTT  
 GGCTTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGC  
 ACAGCCCTATTCTGGCTTGACTACGTTACATGTTGGGTACAATTGCT  
 GAGGATATAGTCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGC  
 GGCTGGGTTTTGTCATGTTCTGAGTGAAGGGCTTCCAGGAGGCAAAGT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATAT  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCATGA (SEQ ID NO: 87)

**love-39**

ATGGCTCAGATCAAGGTATATTACGAACCTCGTCACTCTCTCACCAAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTACTGAAATAAGGAGGTTAATGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGAGCGGATCTCGT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCTCCAGTGCCTCCAGTGCCTCCAGAGCTGCGC  
 TAGACATATCCGAGTCGCATTCTCAAATACCTCCCGCAATTCTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCATTGACGAGGCTATTGACACTGACTGCTGGGGCTGTC  
 ATGTGATGGAGGCTTCAGCTGCAAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAGT  
 CTACGGTTGAAAAGCTCCGTGCCACCGATATCGAGCGACATTGCTCGTGCAGGCCAGTGC  
 CGAGAGCTTTGATGACCTGTCGGCGGTGCGCAGGAACCTGGAAGAGAGATCCTCTGGCGTGAC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTCTTACGGCAGTACACTGTTACATATTGAATGTGCGGATTTGGCCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGCATTGCACTCAGCATGAGCGGGAAACCGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGGCCACACTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGAAGGGCTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGCAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:88)

**love-40**

ATGGCTCAGAACAAAGGTATATTACGAACCTCGTCACTCTCTGCCAGTGGAGGGTTACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTTGTGATCGGTGTCATGCACAAAGATCA  
 AATGTACTGAAATAAGGAGGTTACTGGCCGTGCTCCCTGTCAAGCGTTGCCAGCAGGCTGGACTT  
 CGATGTCGATCACAGTGGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGAGCGGATCTCAT  
 CTCTGCTGACCCAGATCCCTGCTGCACATGTCCCGCCTCCAGTGCCTCCACAGAGCTTGC  
 TAGAAGTATCCGAGTCGCATTCTCAAATACCTCCCGCAATTCTGATCCACCGGACAGCTAC  
 GACTGGTCGTGGACCTCGATTGGCACTGACAAGGCTATTGACACTGACTGCTGGGGCTGTC  
 ATGTGATGGAGGCTTCAGCTGCAAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTTGAGT  
 CTACGGTTGAAAAGCTCCGTGCCACCGGTATCGAGCGACATTACTCGTGCAGGCCAGTGC  
 CGAGAGCTTTGATGACCTGTCGGCGGTGCGCAGGAACCTGGAAGAGAGATCCTCTGGCGTGAC  
 GGTAGAATGGCGAAGCAGGAAATCTGGACCCATCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGGATGTGCGGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAG  
 GGAGTCGATCCCAGTCGCCAGCAGAGACACCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAG  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGATTGAGA  
 TTACTCTGGAGTACACTCCGCCGGGCATTGCACTGCGCTTCCATCAGCATGAGCGGGAAACCGGC  
 GAGGATATAGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGGCCACACTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGAAGGGACTTCCAGGAGGCAAAGTCTGCTGGTT  
 CCCGAGGTGCAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTGCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCATGA (SEQ ID NO:89)

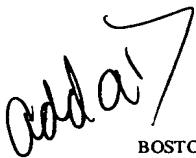
love-41

ATGGCTCAGATCAAGGTATATTACGAACCTGGTCACCTCTCGCCAGTGGAGGGTCACGCAC  
 CGGTGGAACATTACCCGCCGTGCATTCCGACGCTCTTGTATCGGTGTCATGCACAAAAGATCA  
 AATGTAAGGAAATAAGGAGGTTACTGGCGTCTCCCTGTCAGCGTTGCCAGCAGGCTGGACTT  
 CGATGCGTCTACAGTGAGCGATGCCCAAGCGCAAGCTACGCCAATCCAGGGCAGCGGATCTCGT  
 CTCGCTGACCCAGATCCCTGCTGCACATGTCCACCTCCAGTGCCTCACAGAGCTGCCGC  
 TAGACGTATCCGAGTCGCATTCCCAAATACCTCCCGCAATTCTTGATCCACCGGACAGCTAC  
 AACTGGTTGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAAT  
 ATGTGATGGAGGCTTCAGCTGTCAGTTAGAGCCAACGCTGCCGGATCTACCTCGCCCTCGAAT  
 CTACGGTTAAAAAGCTCCGTTGCCACCCGTATCGAGCGACATTGCTCGTCCGCCAGTGCCTGAA  
 CGAGAGCTTTCGATGACCTGTCGGCGGTGTCGCAGGAACCTGGAAAGAGATCCTCTGGCGTGAC  
 GGTAGAATGCCGAAGCAGGAATCTGGACCCATCCCATCGGAATGTTTCAATGCGTCACGAC  
 GGCTTCTTACTGTCCCTGCGCCAACAAGCGCAGGCCACTGCCATCAAGGCACACTAGACGAATGT  
 TTACGGACCAAGAACCTTTACGGCAGTACACTGTTACATATTGAATGTGCCGATTTGACCGC  
 CATATCGGAGTTGCTCCTGTCGCAAATTAGGCGGACCCAGAACAGCCATATGAGCCACTGGAAAG  
 GGAGTCGATCCCAGTCGCCGAGCGGAGACGACACCAGCAGCAGCAGCGGCCACAGCAGTGTGAC  
 ACCATACCCCTTCTTAGCGAGAACCTCCCTATTGGTGAGCTGTTCTCTATGTTGACCCCTGAC  
 ACACGCCCTATTCTGGCTTGCACTACGTTACATGTTGGGTACAATTGCTGCGTGAGAATGAGA  
 TTACTCTGGAGTACACTCCGCCAGGGTATTGCACTGCTCCATCAGCATGAGCGGGAACCAGGC  
 GAGGATATGCCAGGACAGGGCGACCAATTCCGCAAGATGCGAGGAGCAGCCGACCACTCCAGC  
 GGCTGGGTTTGTTCATGTTCTGAGTGAAGGGGCTTCCAGGAGGGAAAGTCTGCTGGTT  
 CCCGAGGTGAAACCATCGCAGCACTGCGACGATGCTATGAGGATATCTTCCCTCGCCCGCAA  
 CACAAACATGGCATGCTCAGAGACCTCAACAATATTCCCTCCATGA (SEQ ID NO:90)

5

## Equivalents

Those skilled in the art will recognize, or be able  
 to ascertain, using no more than routine experimentation,  
 10 many equivalents to the specific embodiments of the  
 invention described herein. Such equivalents are  
 intended to be encompassed by the following claims.



BOSTON 1322173v1